IPME Evaluation – Final Report

October 2017

Submitted by
WYG International Limited

creative minds safe hands

CONTENTS

ABBREVIATIONS AND ACRONYMS

6MR	Six Monthly Report
ASWA	Accelerating Sanitation and Water for All
CATS	Community Approaches to Total Sanitation
CBM	Community Based Management
CBO	Community Based Organisation
CLTS	Community-Led Total Sanitation
CO	Country Office
CS	Customised Support
CSO	Civil Society Organisation
DBE	Department of Basic Education (Myanmar)
DCT	Direct Cash Transfer
DFID	Department for International Development
DHS	Demographic and Health Survey
DPP	Development Program Proposal
DPH	Department of Public Health (Myanmar)
DPHE	Department of Public Health Engineering (Bangladesh)
DR	Document Review
DRD	Department of Rural Development (Myanmar)
DWSS	Department of Water Supply and Sewerage (Nepal)
EAPRO	East Africa and Pacific Region Office (UNICEF)
EQ	Evaluation Question
ESARO	Eastern and Southern Africa Regional Office (UNICEF)
GIS	Geographic Information System
GRPR	Global Review of Programme Results
HACT	Harmonised Approach to Cash Transfers
HAWQ	Hardware and Water Quality
HIV/AIDS	Human Immunodeficiency Virus Infection and Acquired Immune Deficiency Syndrome
HQ	Headquarters
HWWS	Hand Washing with Soap
IEC	Information Education Communications
INGO	International NGO
IP	Implementing Partner
IPME	Independent Process Monitoring and Evaluation
JMP	Joint Monitoring Programme
JRG	Joint Reference Group
KAP	Knowledge Attitudes Practice
KII	Key Informant Interview
LTA	Long-Term Agreement
MDG	Millennium Development Goal
MENA	Middle East and North Africa Regional Office (UNICEF)
MICS	Multiple Indicator Cluster Survey
MIS	Management Information System
NGO	Non-Governmental Organisation
NMIP	National Management Information Project (Nepal)
NUST	University of Science and Technology (Pakistan)
NYHQ	New York Headquarters

OD	Open Defecation
ODF	Open Defecation Free
PATS	Pakistan Approach to Total Sanitation
PCA	Project Cooperation Agreement
PPP	Public Private Partnership
PVA	Programmatic VFM Assessment
OECD	Organisation for Economic Cooperation and Development
OECD-DAC	OECD Development Assistance Committee
RAF	Results Allocation Framework
RAG	Red Amber Green
RO	Regional Office
ROA	Rapid Outcome Assessment
ROSA	South Asia Regional Office (UNICEF)
RWSN	Rural Water Supply Network
SA	Sustainability Assessment
SACOSAN	South Asian Conference on Sanitation
SAM	Sustainability Assessment Myanmar
SC	Sustainability Check
SDG	Sustainable Development Goal
SDP	Sector Development Plan
SEIU	Sector Efficiency Improvement Unit (Nepal)
SHEWA-B	Sanitation, Hygiene Education and Water supply in Bangladesh
SMS	Short Messaging Service
SWA	Sanitation and Water for All
TOC	Theory of Change
TOCR	Theory of Change Review
TOR	Terms of Reference
TPFM	Third Party Field Monitoring
UNEG	United Nations Evaluation Group
UNICEF	United Nations International Children's Emergency Fund
UPENSONG	University of Pennsylvania Social Norm Group (Pakistan)
USD	United States Dollars
VFM	Value for Money
WASH	Water and Sanitation for Health
WASHBAT	WASH Bottleneck Analysis Tool
WCARO	West and Central Africa Regional Office (UNICEF)
WIMS	National WASH Information System
WSP	Water and Sanitation Program
WYG	WYG International

ACKNOWLEDGEMENTS

IPME thanks the Republic of the Union of Myanmar, the Islamic Republic of Pakistan, and the Republic of Madagascar for government support in conducting the evaluation missions. IPME thanks the staff of the UNICEF Myanmar Country Office, the UNICEF Pakistan Country Office, the UNICEF Madagascar Country Office, and UNICEF Head Quarters for their collaboration and support. We also thank all the stakeholders who participated in interviews and workshops and gave generously of their time and experience.

EXECUTIVE SUMMARY

ES1 Introduction

This Final Evaluation Report of the Accelerating Sanitation and Water for All (ASWA) Programme in Neglected, Off-Track Countries (submitted by WYG International [WYG] in association with Aguaconsult) presents findings and conclusions gained through implementation of the evaluation design as set out in the Independent Process Monitoring and Evaluation (IPME) ASWA Programme Evaluation Design Document (see Annex S). This report is presented to DFID – as funder of ASWA – and to UNICEF.

The purpose of the evaluation was to investigate the reasons behind the achievement/non-achievement of verified results of the ASWA program and to gain a deeper understanding on the functioning of selected ASWA country programmes. The objectives of the evaluation, in order of priority for DFID, are to understand and share lessons on:

- The prospects for sustainability of outcomes.
- The extent of value for money.
- How and why verified results were achieved / not achieved.
- The quality of outputs.

The subject of the evaluation is the ASWA Programme at country level (including support from UNICEF Head Quarters [HQ] and Regional Offices [ROs]) between November 2013 and September 2016.

The remainder of this Executive Summary covers:

- **Findings and Conclusions**: on Results Achievement, Quality of Outputs, Outcomes and Sustainability, Value for Money, Equity, Programme Monitoring, and Other Changes and Innovations.

- **Recommendations:** for DFID, UNICEF, and the Water and Sanitation for Health (WASH) Sector.

ES2 Results to Date

For Outputs 1-4 (which relate to the delivery of water, sanitation and hygiene outputs at community level) most country programmes have met or exceeded their output targets, and by a considerable margin in some cases. Where there are shortfalls, these are mostly quite small. Results under Output 5 (enabling environment) are all qualitative and in the absence of country-specific objectives or targets (except in Bangladesh) it is very difficult to confirm to what extent the output has been met.

ES3 Findings and Conclusions

ES3.1 Results Achievement (EQ 2a, 2b, 3, 3a, 3b, 4, 4a, 4b, 7c, 11)

In examining variations in programme results it is important first to acknowledge that there were important differences between the country programmes. This was evident in terms of: the presence or absence of emergencies, conflict or security constraints; sector status in terms of policy, institutional arrangements, technical capacity and external support; country programme design in terms of scope and scale (Nepal had a sanitation target 28 times greater than that in South Sudan); the range and capacity of implementing partners; the availability of tested implementation approaches; and country programme management in terms of logistical challenges and human resource constraints. Against this backdrop, findings on technical components included the following:

ES3.1.1 Sanitation and Hygiene

Outputs: The South Asian country programmes all set ambitious targets, reflecting their long experience of Community Led Total Sanitation (CLTS), its adoption by government as a national approach and a relatively stable operating context. Nepal has a rural sanitation movement with strong government leadership at district level, especially in the Terai region where much of ASWA activity takes place. UNICEF is one of several government and external agencies engaged in a concerted, co-ordinated effort to achieve district-wide Open Defecation Free (ODF) status.

Madagascar was something of an outlier in that it also adopted very ambitious targets despite a weak sector framework, extreme poverty and acute water supply problems. Here enabling factors included an operational model that was tested and improved before ASWA began, and featured the direct deployment of UNICEF consultants into government agencies.

In Cambodia, both UNICEF and government partners had been working with CLTS for some time but programme ambitions were quite modest due to human resource constraints and the fact that operations were spread quite thinly over a wide area, which was challenging for supervision and monitoring.

The Myanmar country office and partners had reasonably limited prior experience with CLTS and earlier projects had not operated at scale. The Niger, South Sudan and Yemen teams were somewhere in the middle in terms of experience; operations in South Sudan and Yemen were also constrained by conflict.

Several COs cited the limited technical capacity of implementing partners as a constraint, nevertheless targets were mostly exceeded. Another constraint cited in some countries was that other sanitation projects operating in locations near to ASWA interventions were providing hardware subsidies, which undermined efforts to generate demand without subsidies.

Defining and measuring people 'reached' with hygiene promotion was problematic, nevertheless interventions were generally substantive and had a strong focus on hand washing in line with indicator 3.2. Enabling factors for hygiene outputs included:

- Including a hand washing facility in the ODF criteria applied in some countries.
- The ready availability in most locations of water for hand washing (which need not be potable), though there were seasonal shortages in some places.
- Promotion of very simple hand washing technology such as the 'tippy-tap' as used in Madagascar, which requires nothing more than an empty plastic bottle, some string and a few sticks, which can be periodically repaired or replaced by the household.
- The acceptance of ash (which is invariably available) as an alternative to soap where this is expensive or not locally available.

WASH in schools was a relatively small programme component overall. An important enabler was the application of UNICEF's 'Three Star' Approach which provided a standard framework for hygiene promotion and the improvement of school facilities. In addition, most COs engaged with national government to pursue the introduction of policies and strategies to bring all schools up to an acceptable standard.

Outcomes: Principal constraints on the potential achievement of outcomes:

- Once COs had received funds and procured Implementing Partners (IPs), the implementation period remaining was much shorter than programme documents suggested and in most cases the Project Cooperation Agreements were only long enough for the delivery of outputs.

- Since IPs are normally only contracted for the programme implementation phase, post-ODF community support and monitoring needs to be provided by government agencies with a permanent presence in the local area. In the three case study countries, UNICEF was engaged in advocacy and technical support to help local agencies take on these responsibilities, but this was a work in progress. This is in fact a long-term challenge for the sector. UNICEF and other development partners ensure that it features in sector debate at policy level as part of wider initiatives to foster sustainability.
- Many of the toilets built in response to CLTS triggering were simple constructions without durable superstructures. Some country programmes including Nepal, Bangladesh and Pakistan have (or have previously had) sanitation marketing initiatives that promote and enable the construction of durable and hygienic latrines in rural areas. However, in sparsely populated rural areas such as southern Madagascar the scope for such market-based approaches is very limited. In such locations, toilets built with locally available materials can function reasonably well. However, they may not withstand heavy winds or rains and will therefore need to be repaired or replaced seasonally. This is not a problem in principle, but may not happen without some ongoing motivational stimulus from government or external agencies. UNICEF COs are trying to address it through advocacy and technical support. However, the challenge is huge as development agencies have little control over government actions post-implementation.

Enabling factors included:

- Where COs applied promotional approaches that had been tested and improved prior to ASWA the prospects for behaviour change were likely to be better than in countries where the approach was newer and needed fine-tuning.
- Where there was a strong, government-led initiative to promote rural sanitation and hygiene, outcomes and sustainability were less dependent on the direct contributions of UNICEF or other external agencies. Nepal stands out as an example of good practice in this regard which is to be commended.

ES3.1.2 Water Supply

Outputs: Most COs (COs) applied established approaches to developing, rehabilitating or improving existing community water points (typically boreholes with handpumps) which were then operated under a community management model. In a few cases, however, small piped schemes were also developed. In Myanmar and Madagascar there was only limited prior experience with the chosen service delivery model prior to ASWA. Also, in Bangladesh the conversion of illegal water connections to legal, metered connections providing safe water to groups of households was expanded under ASWA (building on an earlier programme not led by UNICEF).

Outcomes: The issues and challenges in ensuring that local government agencies take on long term responsibility for community support apply to water supply as well as sanitation. Sustainability depends less on securing behaviour change, but there is the additional challenge of trying to ensure a ready local supply of essential spare parts via the private sector, which can be difficult in remote and sparsely populated areas.

UNICEF has facilitated the involvement of private contractors in managing rural piped water supply schemes in Madagascar. The long-term viability of the service delivery model is still being determined, and in the South, it has been very difficult to attract operators. A more enabling environment is needed in both Myanmar and Madagascar to address barriers to increased private sector participation, particularly in the areas of regulation, procurement and governments' own capacity to engage in public-private partnerships.

The evaluation missions provided only limited exposure to community water points, hence it is difficult to comment further on the prospects for sustainable outcomes. The piped schemes in Myanmar and Madagascar were relatively new initiatives. They were in effect pilots and there was much to be learned about their cost-effectiveness and long-term viability. In Madagascar, UNICEF was closely monitoring the use and performance of the schemes to learn to what extent they were delivering real benefits and what could be done to enhance their sustainability. In Myanmar, however, there was less evidence of UNICEF doing this. There was evidence in both countries that the revenue generated might not be sufficient to cover operation and maintenance costs.

In several countries, UNICEF is helping government tackle sustainability as a strategic issue at national level. The absence of fully functional sector monitoring systems providing reliable data on the quality and reliability of services is a common constraint. In some countries (Cambodia, Nepal and Bangladesh, for example) UNICEF is helping government to establish more useful management information systems. Furthermore, in Madagascar and Pakistan, UNICEF has spearheaded the introduction of periodic sustainability checks to inform sector planning and (hopefully) resource allocation. UNICEF headquarters is promoting and supporting the introduction of sustainability checks by all country programmes.

ES3.1.3 Enabling Environment

Sector Monitoring: There is insufficient evidence available to assess the extent to which sector monitoring has improved in the ASWA countries, because no benchmarks or targets were set at country level as a point of reference. However, in those countries where the IPME team has had some exposure to UNICEF work in this area (this excludes Yemen and Niger) it is evident that the support has been strategically significant, well-targeted and highly collaborative. While there has been some progress, there is a long way to go in most countries. Establishing viable sector monitoring systems is a long-term initiative that rarely fits into the time frame of a single donor-funded project.

Measurement of Indirect Beneficiaries: The initiative to develop a methodology stalled and there was no piloting under ASWA. UNICEF is still pursuing the challenge at global level. We acknowledge the potential benefits of a viable methodology, but question the value of including specific numbers of indirect beneficiaries in future logframes before that methodology has been established.

Strengthening Government-led Scale-up: UNICEF is helping government to scale up WASH interventions (particularly in sanitation) in several countries (e.g. Madagascar, Cambodia, Niger, Nepal). Moreover, in most countries UNICEF is an active participant in sector fora at national level and proactively supports the strengthening of sector planning and co-ordination processes.

Capacity Gaps in Lead Sector Institutions: UNICEF support has been substantial and appreciated by government partners. Having said this, COs are not, in most cases, providing their support on the basis of a systematic needs assessment and capacity building plan that identifies and targets 'critical gaps'. This and a lack of country-specific targets make it very difficult to assess the extent this component of the output has been achieved.

Operational Research: The research commissioned by UNICEF Head Quarters (HQ), while relevant to the UNICEF programme globally and the WASH sector in general, has been largely disconnected from the rest of the ASWA programme. The evaluation has seen no evidence of the emerging findings being used to inform programme strategies in the nine ASWA countries – with the possible exception of the social norms research in Pakistan. The justification for including centrally-managed research in ASWA is therefore unclear. For the ASWA COs, items listed reflect a general lack of clarity in the programme as to whether they should be reporting anything against this indicator; and what counted as operational research rather than operational analysis. Some of what has been reported is clearly not

operational research. Nevertheless, from IPME engagement with COs and reports shared by UNICEF it is evident that some potentially useful operational research has been undertaken.

ES3.2 Quality of Outputs (EQ1,1a,1b, EQ 6a, 6d, 6e)

ES3.2.1 Programme Relevance

Each ASWA country programme was designed as a sub-set of the broader UNICEF country WASH programme, which was itself designed in collaboration with national government and supportive of national WASH policy and/or strategy (if any existed). At the design stage, UNICEF HQ required each country office to show in their proposal how the planned ASWA interventions were aligned with the Intermediate Results set out in the Country Programme Action Plan. ASWA was therefore directly relevant to UNICEF country strategies and, by implication, government strategy and plans.

UNICEF collaborates closely with both national governments and other international development agencies. Under Output 5 (Enabling Environment), many of the country programmes have supported initiatives to strengthen sector policy, strategy, planning, co-ordination or monitoring. Regarding relevance to DFID, Bangladesh was the only ASWA country where DFID had an ongoing bilateral programme with a WASH component, though some others received WASH humanitarian support. Alignment of ASWA with DFID country strategies was, therefore, not a concern.

ES3.2.2 Output Quality

The evaluation missions provided only limited opportunities to assess the quality of physical outputs, whether hard or soft. The water supply and sanitation facilities seen during field visits were generally of a good or reasonable quality, but the number was too small to draw any conclusions for the programme overall. The monitoring systems appraisals carried out by IPME for seven of the nine country programmes in 2015 examined whether programmes featured *"mechanisms to monitor the quality of outputs, both 'hard' (e.g. water points) and 'soft' (e.g. hygiene promotion) and to address any shortcomings identified."* Five out of seven country programmes monitored output quality comprehensively, while two out of seven did it partially. Four out of seven monitored 'hard' and 'soft' outputs separately, while three out of seven partially met the criteria. Furthermore, most UNICEF COs had regular meetings with implementing partners to review progress and address any shortcomings identified. The most comprehensive quality assurance system was found in Pakistan, which had included independent third party field monitoring.

ES3.2.3 Other

ASWA design at country level was in most cases not informed by needs identification of specific target populations. However, some of the most under-served districts and communities were targeted at the planning stage. Thereafter, however, most UNICEF CO's did not track whether they were effectively serving those most in need within these areas.

In 2014, UNICEF HQ issued a WASH Climate Resilient Development Strategic Framework. At the ASWA global meeting in April 2015, UNICEF HQ presented some outline guidance on how COs could conduct 'climate change and water resource assessments' under Output 2. Further guidance was issued to COs, but requiring a new skillset, additional resources, and time. Few countries have conducted such assessments (one very recent exception being work in progress in Bangladesh), in fact this output has effectively been shelved by the COs.

There is insufficient evidence for the ASWA programme as a whole to state how appropriate locations for water supply services were. Uncorroborated evidence from two sustainability assessments (Myanmar, Pakistan) suggest that the majority of water supplies were appropriately located from a water safety perspective. In the locations visited for the evaluation the siting seemed appropriate.

Community Based Organisations (CBOs) with responsibility to operate and maintain water supply schemes require further external support to address capacity gaps in technical and management skills. The respective roles and responsibilities of CBOs and government agencies also need clarification, especially regarding major repairs.

ES3.3 Outcomes and Sustainability (EQ 2, 6, 6b, 6c, 7, 7a, 7b, 7d, 10, 12)

ES3.3.1 Outcome Data

Virtually no relevant outcome data (endline versus baseline) was available for the majority of ASWA programme countries by 30 June 2016 (the agreed deadline for evaluation evidence). By the end of 2016 some relevant surveys had been completed, but their usefulness to the evaluation was constrained by a number of factors. In particular: sustainability checks in Pakistan and Madagascar were sector-wide and not closely aligned with ASWA logframe indicators; the Pakistan endline survey only gathered output (access) data; and the Bangladesh and Myanmar endline surveys had data quality issues. Consequently, it is only possible to analyse prospects for sustainable outcomes and draw findings and conclusions in general terms that broadly reflect sector conditions and may not fully reflect the ASWA programme.

Contributory factors to this lack of outcome data include: outcome targets not being set at country level; the logframe assumption of 100% conversion of outputs to outcomes; a lack of clarity at programme start on what COs should do on outcome assessment; and the drive to achieve output targets. The wider IPME team consulted COs about their plans to produce outcome data in early 2015. It also made a presentation and an offer of technical support on collecting outcome data at the ASWA Global Meeting in March 2016. Two countries took up this offer (Pakistan and Bangladesh) which contributed to the development of sustainability checks.

ES3.3.2 Sanitation

Some people are using sanitation facilities, but (as is already recognised in the WASH sector globally) increasing access to latrines by itself is not eliminating Open Defecation (OD). Elimination also requires behaviour change (as planned by ASWA) that is sustained over time to adopt sole use of latrines for defecation. Communities are willing to participate in and mobilise for ODF when CLTS methods are adapted based on evidence from national implementation that reflects contextual factors, including social norms and village size. However, as is already known in the sector globally, the co-existence of subsidy approaches can be a constraint to community willingness to participate in and mobilise for ODF.

Communities' willingness to sustain ODF has been improving, but rapid ODF slippage is a real risk if there is a gap between the departure of implementing partners and the take-up of responsibilities for implementation of post-ODF activities by responsible government and community actors. This can be because these actors require additional funding (or need to prioritise existing funding) for these roles.

The choice and implementation of strategies for the removal of barriers to the installation and sustained use of improved latrines need to be based on evidence about national contexts, including social norms and baseline levels of latrine access and OD. It should not be assumed that CLTS is appropriate in all national or sub-national contexts. For example, in national contexts where there is evidence of existing higher levels of latrine access, sanitation marketing (with market / supply chain enabling environment support as needed) may be a valid approach in combination with or instead of CLTS.

Sanitation monitoring and reporting systems in the WASH sector in ASWA countries need strengthening, including verification systems.

ES3.3.3 Hygiene

Few people are adopting Hand Washing with Soap (HWWS) after defecation. There is some evidence of increases over baseline (of between 9-20% from baseline) in household investment in facilities that would enable them to practice HWWS (which is in the middle range compared to evidence in other WASH evaluations reviewed which showed 13-46%). Although this is not a direct indicator of handwashing practice (but is often used as a proxy), such an increase is a positive result. There are outstanding needs for HWWS awareness raising and follow-up, and a lack of clarity of how this will be funded going forward.

ES3.3.4 School WASH

Due to insufficient evidence, there are no findings and conclusions on the outcomes and prospects for sustainability of school WASH for the ASWA programme as a whole. This is because baselines and special studies to assess ASWA outcome indicators for school WASH have not been planned and / or completed by UNICEF for all nine ASWA countries. In addition, for the two case study countries that included the school WASH (Pakistan and Madagascar), the sector sustainability assessments did not cover school WASH. Indicators for outcomes and prospects for sustainability of school WASH should be included in the design of future baselines, special studies to assess outcomes, and sustainability assessments.

ES3.3.5 Water Supply

Choice may be a factor in continued use of unsafe water sources some of the time even when an improved source is available. This may be for reasons of taste or a choice to only use safe water sources at critical times, but further evidence is needed. Water systems may not be financially viable because the costs of operation and maintenance are not fully covered by tariff levels and revenues, which may themselves be constrained by willingness and/or ability to pay.

ES3.3.6 Enabling Environment

Funding to sustainably support WASH sector systems (both institutional arrangements and infrastructure) beyond external programme funding does not appear to be in place. WASH ministries have ambitions and plans to increase WASH finance but it will require strong advocacy across government and with development partners to fill funding gaps. Government support to decentralised administrative systems is partially adequate, with needs for clarity in roles and responsibilities and their financing. At present these can reflect silos and/or overlap between separate ministries that have stakes in WASH at national level. UNICEF investments in learning and coordination are valued by government and other development partners and it is viewed as a leading contributor to such efforts (see also ES 3.1.3 Capacity Gaps in Lead Sector Institutions, ES 4.1.5 Rec 17, and ES 4.2.1 Rec 5).

ES3.3.7 Overall

Most UNICEF COs implementing ASWA understand the need for support to WASH benefits after outputs have been delivered and plan in collaboration with communities and government for this. However, many COs also have remaining concerns about sustainability despite these plans. It is not clear that sufficient resources have been dedicated by ASWA within the implementation phase to support beneficiaries going forward. From a Value for Money (VFM) perspective, funds explicitly dedicated to post-output support, and carefully expended have potential to increase sustained outputs.

UNICEF has been encouraging COs to develop plans for sustainability assessments. By the end of 2015, only Madagascar and Pakistan had concrete country-specific sustainability plans. Only five of nine ASWA countries requested IPME support for sustainability planning through the first quarter of 2016 (South Sudan, Nepal, Madagascar, Pakistan, and Myanmar).

Regarding the overall validity of this evaluation's findings, no strong evidence has been found that would constitute a plausible 'direct rival' explanation for ASWA programme level outcomes (i.e. where other development partners' programmes account for the results). WASH and/or humanitarian programmes (with WASH inputs), at the same scale as ASWA with related outputs and outcomes, are not being delivered by other development partners in many of the same locations as ASWA. Further evidence would be needed to corroborate or reject whether a 'commingled rival' explanation for any ASWA programme level outcomes is plausible (i.e. where ASWA and other development partners both contributed to the results).

Insecurity, conflict, political instability, elections, natural disasters or epidemics are not reported to have displaced people in ASWA locations (even when wider populations nationally may have been). No strong evidence has been found for natural, economic forces, or pandemic disease having led to the displacement of target populations that would constitute a plausible 'super rival' explanation for ASWA programme level outcomes (i.e. where a force majeure larger than ASWA accounts for the results).

ES3.4 Value for Money (EQ 5, 5a, 5b)

ES3.4.1 Effectiveness and Equity

VFM) in terms of effectiveness cannot be fully analysed as valid outcome level data was not available for the majority of ASWA countries or at programme level. Qualitative assessment of the potential for improved outcomes shows that programme level achievement has been above targets for key output indicators. This is a positive indicator of programme traction, and sets a strong foundation for sustained outcomes and effectiveness. To strengthen effectiveness, DFID and UNICEF should allocate a proportion of funds specifically for follow-up to support the transition from outputs to outcomes within the funded implementation phase.

VFM in terms of equity cannot be fully assessed as gender disaggregation is the only measure for which data are available. Moreover, UNICEF gender disaggregated data is estimated and largely follows general population trends. Thus, variations in access to and use of benefits are assumed and these estimates serve little purpose. Equity measures beyond gender estimates should be undertaken, especially where beneficiaries may be denied access to benefits due to caste, ethnicity, cultural barriers, or wealth. In a programme where considerable programme flexibility to shift funds across outputs and countries exists, it is imperative to measure equity of access to ensure that programme flexibility does not undermine Output equity. Some implementing partners in some countries do in fact have disaggregated household data, but have not always reported it if UNICEF did not request such data.

In planning and reporting, both UNICEF and DFID should be aware of the potential negative impact on effectiveness and equity. This may be caused by the drive to maximise output results and the ongoing unmet burden of funding and support to sustain WASH outputs after ASWA ends (see also ES 3.5 Equity below).

ES3.4.2 Economy and Efficiency

The programme is delivering VFM in terms of economy and efficiency, with some variation across countries. Average ASWA programme unit costs are below those budgeted in the 2014 ASWA Business Case for three of four key outputs. Indirect to direct costs are below 10% in four of nine ASWA countries (Bangladesh, Madagascar, Niger, and Pakistan); and below 15% in seven of nine countries (Myanmar, Yemen and South Sudan are between 10% and 15%). Nepal and Cambodia are higher (20% and 31% respectively). For the programme as a whole indirect to direct costs are under 11%, demonstrating an apparent economy of scale provided by UNICEF.

The fact that CO procurement and budget management is embedded in global UNICEF systems enables timely financial analysis and clear metrics and procedures for economic procurement. Most CO

procurement for ASWA is by IPs, which greatly increases the efficiency and timeliness of output delivery. Madagascar, Bangladesh, Nepal, and Pakistan employ well established protocols within UNICEF to achieve procurement of economic and results-oriented partners. Madagascar's use of Long-Term Agreements to prequalify a pool of suppliers, then hold the suppliers to a fixed price contract payable on results is a potential lesson to be shared across ASWA country programmes (see Annex P, p.27). Also, worth mentioning is the Results Allocation Framework (RAF) methodology used by the Nepal CO. This is an effective method to direct and budget funds through partners with agility for results (see Annex P, p.27).

Fund reallocation to country programmes with capacity to scale quickly, is a high value VFM approach, maximising both economy and efficiency. However, by reallocating funds for greater results, there is potential that funds will be moved away from difficult-to-reach populations, thus compromising equity which would undermine VFM. Despite this real tension, we view the reallocation of funds to achieve results as powering positive economy and efficiency. Strong and consistent monitoring support, and strategies to strengthen sustainability of WASH inputs are also vital. If the drive for beneficiary numbers incentivises programmes to concentrate on overachieving their output targets at the expense of focusing on longer-term outcomes, the effectiveness and VFM of initial investments could be undermined.

To facilitate improved costing, DFID and UNICEF are encouraged to compile unit costs for key interventions across countries, ascertaining differences in the components for unit-costs so that a range of costs for standardised components by intervention is gradually developed.

ES3.4.3 Evidence Streams

Overall: The conclusions reached about ASWA VFM are supported by review, analysis, disaggregation, and triangulation of quantitative (financials and indicators) and qualitative data (UNICEF country reports, HQ procedures, and DFID approaches) from multiple sources. These are primarily from the country level but complimented by HQ and Regional levels and IPME monitoring and verification data.

Economy and Efficiency VFM: At the country level, output results data appear to be credible for key Outputs, though there was initially some misunderstanding around the meaning of output indicators 1.1 and 1.2 and how to count the results.

The degree to which available financial data are informative for VFM analysis is in question. Specifically, CO budgets appear to be regularly re-adjusted to meet actual expenditures if there is a significant variance from the planned budget. Best-practice financial reporting is to ensure that a record of budget changes is clear so that there is a trail of evidence to understand how and why changes were made. When the budget is adjusted to resemble the actual expenditure, the reasons and context for programme changes remain obscured. There is value in agile budget readjustment and reallocation of resources to other needs. However, budget readjustment to closely match expenditures does not permit VFM analysis to examine the efficiency of initial programme planning and budgeting. The implication is that when budgets are substantially under or overestimated, the efficient use of funds is undermined, as is VFM.

Also at the country level, some indicators are not associated with disaggregated budget and expenditure data, the result of combining different types of activities and outputs (sanitation and hygiene education, for example). This further weakens VFM analysis of specific outputs.

At the programme level, the strengths and weaknesses of CO performance and financial data are transferred upstream. There does not appear to be significant quality assurance of CO data presented to the programme level. It will be useful at the programme level to establish clear protocols for COs to

follow when reporting financial data (budgets and expenditures) and performance data (targets and results). Such protocols should also set out how indicators are defined.

Budget and expenditures for Regional Offices (ROs) are clear, but it is not clear how such funds were used to support ASWA COs. UNICEF standardized frameworks for financial, results, and sustainability data capture and reporting should be designed, agreed upon, and communicated to recipient countries before the next phase of WASH programming commences with UNICEF. This is needed to ensure standardized application by COs when implementing programmes such as ASWA which require aggregation back up to programme level.

Effectiveness and Equity VFM: Programme-level assessment of the effectiveness of ASWA interventions is not possible because valid outcome level data was not available for the majority of ASWA countries. As further ASWA programming is developed and funded, allocations to gather and assess Outcome data from current ASWA programmes would strengthen future VFM analysis. Evidence to support equity measures in ASWA are more limited than desired. Data capture to demonstrate the equity dimension of ASWA is not well-designed relying, largely, upon assumed gender disaggregation in the general population. Additional equity measures (wealth quintile, caste, under 5 beneficiary) would be useful to demonstrate equity in future programming.

ES3.4.4 VFM Capability

The majority of UNICEF COs in ASWA have come to appreciate VFM analysis as an aid to making informed programme decisions. The challenge and opportunity is to identify and customize VFM support for each CO. The needs of each are different, but the majority are poised and eager to take VFM forward. One constraint to more use of VFM by COs is the need to fund VFM support.

ES3.5 Equity (EQ 8, 8a, 8b)

ES3.5.1 Targeting of Benefits

The ASWA logframe from DFID did not provide a quantitative target to make explicit what inclusion of beneficiaries in the lowest wealth quintile for water services and sanitation facilities would look like. Neither did it include any type of indicator for gender. Greater clarity from the outset would have helped to drive equity planning and outcomes. In terms of ASWA design, only a partially appropriate approach to targeting benefits at communities in the lowest wealth quintile was used by UNICEF. Partial because the approach only relates to water or sanitation benefits (when the logframe suggests it should relate to both), and because the approach is insensitive, as to whether it over or under includes the whole of the lowest wealth quintile. UNICEF's approach to targeting benefits at communities by wealth presents two design risks. Firstly, the risk of inclusion of wealthier households / communities and exclusion of poorer households / communities. Secondly, the risk of implementation errors not being systematically managed or lessons learnt due to no formal monitoring of equity outcomes. A design to benefit the whole community doesn't necessarily mean outcomes are equitable unless targeting, baselines and monitoring of equity is used during implementation.

ES3.5.2 Approach to Promoting Equity

UNICEF's approach to achieve equity through community wide / universal benefits for WASH within geographically selected areas, without specific targeting by wealth quintiles or at women and girls, has been operationalised. Within communities it does use subsidies to target water or sanitation benefits at the poorest 5% or poorest 20 households, in some countries. There is generally a lack of information from UNICEF on the extent to which its approach to promoting equity has been operationalised and successful or not. This is because reporting on equity has not been mainstreamed in ASWA.

ES3.5.3 Processes and Tools

UNICEF has guidance / plans and good practice / insights on gender and WASH (e.g. on identification, monitoring, and evaluation) which were not applied in ASWA and could have driven better equity planning and outcomes. In many countries processes and tools are not in place under ASWA to identify, target and monitor equity between population groups with different social characteristics based on their needs. In a few countries, UNICEF does monitor aspects of equity, but even here this does not appear to be for all variables or ASWA locations. UNICEF HQ / RO has supported countries to better address and report on gender issues. However, more needs to be done in this area. This includes support to COs in developing capacity in UNICEF good practice in equity identification, targeting and monitoring relevant to WASH. Equity guidance to sanitation activities that aim to deliver improvements across entire communities using CLTS methods may require further research in the WASH sector.

ES3.6 Programme Monitoring

ES3.6.1 Output Level

Most country programmes supplied IPs with proformas for monthly or quarterly reporting but provided only limited guidance on how the source data should be generated and managed. Some IPs and UNICEF COs faced human resource constraints, limiting their ability to deploy full-time monitoring staff. Several country programmes commissioned independent baseline studies but these were of limited value for output monitoring when based on limited samples. Community profiles produced in each village as part of the CATS process were far more useful to UNICEF and IPs.

The weakest area of output monitoring was counting the number of people 'reached' by hygiene promotion (and avoiding double counting when the same people were reached through multiple activities). Monitoring by government agencies of IPs' or their own UNICEF-funded activities was generally weak. COs recognised the need to promote, enable and monitor the sustainability of programme results, but few had incorporated sustainability factors into routine monitoring. This said, an increasing number were supporting the introduction of sector-wide sustainability checks. On the establishment of post-ODF monitoring and support, this remained a work in progress in most cases.

COs made use of monitoring information in regular progress review meetings with IPs. However, reporting formats tended to focus mostly on quantitative results with less attention to the quality of processes and outputs.

Alignment with sector monitoring systems was a hypothetical issue in most ASWA countries as none had a system that was operational and effective nationwide (as is typical in many countries globally). Few COs had established mechanisms for linking programme costs and results as part of programme monitoring, though VFM training and technical support from IPME was helping to address this.

ES3.6.2 Outcome Level

Most UNICEF CO's did not expect IPs to report outcome data and this is likely to be accounted for by three factors. Firstly, the programme timeframe was short and, in nearly all countries, field work started late. The focus of programme activity and monitoring was therefore on delivering the outputs which made the outcomes possible. Furthermore, no guidance had been issued on DFID expectations in this area. Secondly, Project Co-operation Agreements (PCAs) tended to be short (typically 12 months) and ended with the delivery of outputs. Thirdly, changes in the number of people using toilets and practising HWWS could not be measured through short visual checks as part of routine reporting; special studies were required.

There was a lack of clarity and consensus between DFID and UNICEF regarding outcome assessment. Irrespective of DFID advice, UNICEF should have made arrangements for assessment of logframe

outcome indicators. However, no relevant outcome data (progress against baseline) was available by 30 June 2016. Two key factors drove this shortcoming. Firstly, at inception, how and when COs would measure progress was not given sufficient consideration by UNICEF HQ and / or COs. Also, very few COs created baselines for ASWA. IPME's appointment was too late for it to influence or support the design of baselines. Though the need to assess outcomes was flagged later by DFID and IPME. Secondly, both DFID and UNICEF were committed to meeting beneficiary output targets. Following a slow start, DFID concerns about progress led UNICEF to accelerate implementation and these targets were eventually met. UNICEF did not intend to marginalise outcome assessment. However, several COs were still focused on output delivery by the time of original target date of March 2016. Inevitably, outcome assessments were pushed back. One other limitation of programme monitoring and reporting was the occasional mismatch between CO data in six monthly reports and the content of global reports from UNICEF HQ. Mathematical errors account for some of this but the reasons for other variances are unclear.

ES3.6.3 IPME Practice

This evaluation did not extend to the IPME team evaluating itself, but it is useful to consider how IPME worked in practice. IPME developed productive working relationships with COs, helped by its demand responsive monitoring Technical Assistance (TA) role in addition to programme monitoring quality assurance. UNICEF COs appreciated the TA and were generally receptive to advice given. The accountability component was significantly affected by starting a year after ASWA. For example, ASWA was half-way through its original implementation period by the time the first monitoring systems appraisals were completed by IPME. Consequently, the time for COs to implement recommendations was limited and often accepted recommendations were scheduled for action in future programming. For example, accepted recommendations in Myanmar were actioned by the time of the evaluation, but those in Bangladesh and Pakistan were not.

Overall, the IPME team feel their engagement was very productive, and informal feedback from UNICEF received outside of the evaluation tends to confirm this. However, there were three limiting factors. Firstly, requests for demand responsive TA typically came from stronger programmes rather than those that arguably needed more support. Secondly, security issues prevented any Yemen visits and limited field visits in Niger and South Sudan. Thirdly, while most IPME recommendations were accepted in principle by COs, they were under no obligation to act on them quickly. In future, IPME (or its equivalent) should be in place at the start of the programme to support and quality assure baseline surveys, and conduct monitoring systems appraisals much earlier. Subsequent output verification should occur at agreed intervals. In addition, to secure timely action on recommendations, periodic tripartite meetings between UNICEF, DFID and IPME (or its equivalent) should review latest IPME reports and agree on any remedial actions.

ES3.7 Other Changes and Innovations (EQ 9, 9a)

There is insufficient evidence from the three case studies or from analysis across all nine ASWA countries to draw findings and conclusions on what other changes (positive/negative, direct/indirect; intended/unintended) have occurred because of ASWA interventions. Uncorroborated anecdotal reports of other positive changes are presented for illustration only. In Madagascar, ASWA school WASH support is reported to be: increasing children's demand for water; inspiring other schools to install hand washing points; and leading children to share hygiene practices with their families. In Pakistan ASWA sanitation support is reported to be: reducing conflict between households because they are not putting waste on each other's land; increasing communities' capacity and mobilisation to access government support on other WASH issues; increasing awareness of disability issues; and increasing school enrolment and attendance. In Myanmar, ASWA water systems support is reported to be: improving school attendance through time saved not collecting water; increasing school access to safe water through connection to new piped water systems; creating tariff funds that can also be used to hire

teachers and repair / improve school buildings; and enabling the expansion / start-up of marketable agricultural produce in some villages.

There are a few instances of innovative approaches being used within ASWA (small piped water schemes in Myanmar and Madagascar, use of mobile phones for monitoring in Madagascar and Bangladesh, and use of social norms methods to extend of CLTS approaches in Madagascar). However, these are not, so far, being replicated beyond the initially intended reach of the programme.

ES4 Recommendations

UNICEF have responded positively to DFID's overt focus on beneficiary numbers at output level and it is commendable that the programme has over-achieved most of its targets in this area. There has, however, been much less attention to outcomes; targets were not set at country level and no common understanding was established between DFID and UNICEF at programme start on how outcomes would be assessed. This, plus the compressed timeframe resulting from the time taken for county offices to receive DFID funds and appoint implementing partners, meant that COs had little or no data to share on outcome level achievements by June 2016.

Similarly, while there is no doubt that equity and sustainability are priorities for DFID, both have been overshadowed to some extent by the focus on output results and again, DFID expectations in these areas were not made clear.

Another limitation of the ASWA programme design was that, while it envisaged large numbers of beneficiaries overall, it did not define a level of ambition at country level in terms of improvement over baseline or the geographical concentration of efforts to achieve a 'critical mass' of promotional effort and capacity development in support of lasting change. Increasingly, sector programmes are targeting district-wide results and adopting strategies for working at scale that leverage the government institutional framework, with strategic interventions at policy level to help create a more enabling environment. The ASWA logframe did refer to supporting at-scale implementation led by government under Output 5, but in practice this was not given much emphasis in the dialogue between DFID and UNICEF.

The Business Case included a Theory of Change (TOC) but this was poorly formulated, incomplete and disconnected from the document text. It did not fully represent the evidence set out the Business Case for linking mechanisms that explained transitions between inputs, outputs, outcomes and impact; did not make explicit where the assumptions in the Business Case fit within these relationships; and was not fully aligned with the ASWA logframe, which was developed after the TOC. Moreover, COs were largely unaware of it and it did not speak fully to their ongoing activities. The original TOC was, therefore, of limited value as a planning or monitoring tool – and DFID had not designed it as such.

A further consideration is that ASWA operated more in the form of nine country projects funded from a common source rather than a fully co-ordinated multi-country programme pursuing a common agenda. There was some facilitated sharing of experiences between countries (not least via two global ASWA meetings and periodic regional WASH meetings[1]) but no ASWA-specific learning strategy or workplan was adopted to ensure that the whole was more than the sum of its parts. Furthermore, ASWA-funded operational research commissioned by UNICEF HQ was conducted in isolation from the ASWA country programmes and has not, it appears, been used to inform programme implementation strategies. To a large extent, this reflects UNICEF's very decentralised management structure and the fact that the ASWA design did not envisage COs doing anything markedly different to their established

[1] A third took place in February 2017, beyond the period under evaluation.

practices. ASWA did not set out to break new ground except, perhaps, in the measurement of indirect beneficiaries and this was dropped from the programme in 2015.

With these points in mind, the following are recommended for any future DFID-funded WASH programme implemented by UNICEF. **Priority recommendations are in BOLD text** and other recommendations in plain text.

Please note that several of the recommendations are very similar to those shared with DFID in July 2016 in response to a request for information on lessons learned from ASWA to inform the design of a follow-on programme.

ES 4.1 Recommendations for DFID

ES4.1.1 Programme Design

1. **The design of future programmes should prioritise, and make specific provision within the implementation phase for, the achievement and measurement of programme outcomes.**

2. The design should also define minimum levels of ambition in terms of improvement over baseline and district-wide impacts, in line with the agenda set by the Sustainable Development Goals.

3. **If another centrally-funded programme is proposed, each participating country office should be required to develop its own logframe or results framework, nested within the global one. This should set out country-specific output and outcome targets, including programme objectives relating to the enabling policy and institutional environment.**

4. Related to 3 above there should be a more overt focus on supporting government-led strategies for working at scale, including country-specific objectives and progress indicators.

5. If a TOC features in a new programme design, it should be sufficiently detailed to be a useful point of reference for programme monitoring and evaluation; aligned with the programme logframe; and integrated with the programme design overall. If Recommendation 4 is adopted, nested country specific TOCs would also be appropriate.

6. If a future programme is to devote sufficient time and resources to the pursuit of outcomes, a two- to three-year implementation period will not be sufficient; five years would be more realistic.

7. Quantitative targets for the inclusion of beneficiaries in the lowest wealth quintile, and for addressing gender-related challenges, should be included in programme logframes to drive equity planning and outcomes.

ES4.1.2 Programme Management

8. Seek clarification from UNICEF on the roles and responsibilities of its global headquarters and regional offices in programme management, implementation, monitoring and learning, given especially the decentralised nature of UNICEF operations.

9. **A dedicated inception phase should be built into programme design. During this phase UNICEF's focus should be on: finalising country-specific logframes / results frameworks and TOC; appointing and training / orienting IPs and / or government**

counterparts; establishing monitoring and reporting systems; conducting baseline studies; and generally ensuring a common understanding among programme stakeholders at all levels of the programme logframe and how results will be pursued and tracked. Improved communication between UNICEF CO, RO and HQ levels will be essential in these tasks.

ES4.1.3 Monitoring and Evaluation

10. DFID should be more proactive during inception to ensure that UNICEF is applying good practice standards and protocols for baseline surveys, output monitoring and outcome assessment (including for school WASH). DFID cannot be too directive, but clarity on DFID expectations in this area would be helpful to UNICEF, and recent experience from ASWA and the WASH Results Programme puts DFID in a strong position to offer advice on good practice.

11. Measuring hygiene promotion inputs (people reached) has proved to be very problematic and arguably adds little value to programme implementation. It is recommended that indicator 3.1 does not feature in future programme logframes.

12. Indicator 3.2 (water resource and climate change assessments) has also proved to be problematic due to a lack of clarity on what needs to be done. DFID should agree with UNICEF what can and should be done in this area, nevertheless the indicator should not feature in future logframes because it relates to an input, not an output.

13. Clarify DFID expectations around the reporting of disaggregated results in terms of gender, ethnicity, socio-economic status and disability at output and outcome level.

14. **DFID should encourage UNICEF to regularise VFM analysis for global and country programmes. Strengthening and regularising VFM assessments will add another dimension to UNICEF's existing monitoring and evaluation structures. To add VFM analysis effectively may involve incremental changes to budgeting granularity, and tracking of expenditures at the activity, output, and outcome levels across interventions and COs.**

ES4.1.4 Sustainability

15. Given some of the recommendations of the recent report from the Independent Commission for Aid Impact (2016) *Assessing DFID's Results in Water, Sanitation and Hygiene, An Impact Review*, DFID should give sustainability a much higher profile within programme design and ensure that targeted post-output resources are allocated accordingly. There is already an initiative within UNICEF globally to ensure that each country office introduces sustainability checks and gives more attention to sustainability within country programmes generally. This remains optional at country level, however, given UNICEF's heavily decentralised institutional structure (only Pakistan and Madagascar have conducted at least one sustainability check and Bangladesh may fund a new sustainability check in future). DFID should encourage and build upon these efforts, supporting UNICEF in addressing sustainability both within its own programme and at national level as a sector-wide issue.

ES4.1.5 Enabling Environment Objectives

16. Enabling environment support is highly relevant, often highly valued by government, and should feature in future programmes. However, objectives need to be customised to specific country situations and tailored to programme timeframes if they are to be both achievable and measurable. At the same time,

programme design should not encourage UNICEF to take on all sector challenges at policy and institutional level, but instead to identify where it can best add value. This should take governance and political economy issues into consideration and the activities of other development partners in the WASH sector nationally. Sector monitoring is a good example: every country needs a viable monitoring system but it should not be assumed that UNICEF will be the lead support agency for this, or that real progress can be made during the programme timeframe. Government may not be open to it.

ES4.1.6 Learning and Research

17. ASWA global meetings have been appreciated by programme staff and partners, but more could be done to facilitate synergies through inter-country and global learning. Future programme designs should include learning objectives and a learning strategy with specific, budgeted outputs.

18. If further operational research is to be supported, there should be a clear rationale for it in it the programme design accompanied by specific objectives for using the findings to strengthen country programmes.

ES4.2 Recommendations for UNICEF

ES4.2.1 Programme Design and Management

1. Complete ongoing work to clarify the roles and responsibilities of UNICEF headquarters and regions in programme management, implementation, monitoring and learning.

2. UNICEF headquarters introduces many WASH-related initiatives at global level but, given the decentralised nature of the organisation, these are not always taken up at country level. For the any future DFID-funded WASH programme, clarify whether and how specific global initiatives will be taken up at country level (for example sustainability checks, VFM analysis) and how these will support the achievement of programme objectives.

3. Set country-level outcome targets and make explicit plans for outcome assessment beginning with baseline assessments that will allow progress against outcome indicators to be measured at specified intervals. If necessary these should extend beyond the duration of the programme implementation phase.

4. Define country-specific objectives and targets to strengthen the enabling policy and institutional environment that are clear, measurable and realistic given the programme timeframe, with a strong emphasis on sustainability. The rationale for UNICEF (rather than other support agencies) taking on these tasks should be explained in country office proposals.

5. At the planning stage, conduct capacity and risk assessments to determine whether IP and / or government counterpart capacity is likely to be a constraint on programme delivery. If it is, then make explicit provision to address this via capacity building support and/or other compensating strategies (such as deploying personnel within partner organisations).

6. Ensure that the duration of IP Project Cooperation Agreements (PCAs) and / or Memoranda of Understanding (MOUs) with government counterparts is sufficient to enable delivery of outputs and preliminary work (at least) in support of outcomes and sustainability.

ES4.2.3 Monitoring and Evaluation

7. Provide detailed guidance to IPs and / or government counterparts not just on reporting but on baseline surveys and how routine monitoring data should be generated and processed. Be clear as to what extent IPs and / or government counterparts will play a role in outcome assessment.

8. The use of third party field monitoring in Pakistan has delivered significant benefits and the use of similar arrangements in other large country programmes warrants consideration.

9. Ensure that all country programme monitoring systems, and associated guidelines, incorporate explicit measures to:

   - Track the establishment of enabling conditions for sustainability in targeted locations
   - Track the extent to which the poorest and most vulnerable members of targeted communities benefit from programme interventions
   - Link programme costs with results (links to recommendation 14 below).

ES4.2.4 Value for Money

10. To strengthen the VFM effectiveness of ASWA country programmes, UNICEF should allocate a proportion of funds specifically for follow-up to support the transition from outputs to outcomes within the funded implementation phase.

11. Greater clarity is needed in reporting the use of, and results gained from, funds expended by the Regional Offices to support COs and their role in doing so.

12. To facilitate improved costing, DFID and UNICEF should compile unit costs for key interventions across countries, ascertaining differences in the components for unit costs so that a range of costs for standardised components by intervention is gradually developed.

13. There is a need for transparent tracking of country office budget re-adjustments made to meet actual expenditures when there is a significant variance from the planned budget.

14. All output indicators at country level should be associated with disaggregated budget and expenditure data to avoid combining different types of activities or outputs (sanitation and hygiene education, for example) and to strengthen the potential for VFM analysis of specific outputs.

15. To support quality assurance of country office data by headquarters, UNICEF standardised programme level frameworks should be established for COs to follow when reporting financial data (budgets and expenditures) and performance data (targets and results) including how indicators are defined.

16. Allocations to gather and assess outcome data from the current ASWA programme should be made to strengthen future VFM analysis of programme-level effectiveness.

ES4.2.5 Outcomes and Sustainability

17. Securing funded and timely arrangements for long-term promotional interventions, technical assistance and monitoring in programme communities post-ODF and after IP PCAs and / or government counterpart MOUs have ended should remain a high priority for country programmes. This is likely to require advocacy and technical

support for government and other responsible actors at both local and policy level and applies to both water supply and sanitation/hygiene.

18. More research is needed in some countries (e.g. Myanmar, Madagascar) to explain why some communities continue to use unsafe water sources even when an improved supply is available, and to identify how this might be resolved.

19. Where new service delivery models have been introduced based on piped water supply with shared or household connections and metering, it is important that their technical and financial viability is closely monitored. This will provide opportunities to maximise learning on whether and how they are (or can become) sustainable, cost-effective and scalable options for serving low-income rural communities. In Madagascar, much is already being done in this area with more planned. In Myanmar, however, little has been done on this area so far. There seems to be an implicit assumption that the service delivery model is viable though this is yet to be proven. For both countries, ensuring adequate revenue to fund operation and maintenance in the context of very poor communities is a critical challenge, as is the technical and managerial capacity of the operators in some parts of Madagascar.

ES4.2.6 Equity

20. Output results data should not be disaggregated based simply on the gender balance in the target population. Gender disaggregated primary data (e.g. gender balance of water / sanitation management committees; extent to which water and sanitation facilities are designed, constructed and managed to reduce the risks of gender based violence for women and girls) needs to be collected for monitoring and reporting. Such data should be routinely requested from IPs, some of whom already collect it, and included in outcome survey designs.

21. Equity approaches that aim to include benefits for people in the lowest wealth quintile, but which do not robustly identify wealth distribution baselines within communities, present a risk. UNICEF equity approaches need to, firstly, further mitigate the risk of inclusion of wealthier households / communities and exclusion of poorer households / communities. Secondly, they need to further mitigate the risk of implementation errors not being systematically managed or lessons learnt due to not formally monitoring equity outcomes.

22. Guidance and capacity building for COs in equity identification, targeting and monitoring relevant to WASH is needed and can build on existing UNICEF guidance / plans and good practice / insights and clarification of DFID expectations recommended in this area (see DFID Rec. 14). In addition, robust process and tools to identify poor, marginalised and vulnerable populations (including women and girls as a specific group) and to monitor the outcome of targeting on them should be implemented. It should not be assumed that better targeting of sector resources to provide WASH services to poor, marginalised and vulnerable populations and VFM equity can be achieved otherwise.

ES4.2.7 Research and Learning

23. Future programme designs should include learning objectives and a learning strategy with specific, budgeted outputs.

24. At the design stage, clarify how any proposed research will support the achievement of programme objectives.

ES4.3 Recommendations for the WASH Sector

While the evaluation has not generated immediate recommendations for the WASH sector, there is potential for further learning from ASWA that could be of value to the sector. In particular, this learning relates to the following areas:

1. *The use of Third Party Field Monitoring (TPFM) in Pakistan.* This is not the first UNICEF country programme to use TPFM, but in this case the system has been fine-tuned over several years and offers useful insights to programmes considering the adoption of something similar. IPME and the Pakistan CO have already developed a short Field Note on the Pakistan experience but a more detailed study might be useful.

2. ***The service delivery models adopted for rural piped water supply in Myanmar and Madagascar*. These models are still relatively new, and somewhat unusual, and their viability as options for serving the rural poor at scale in the long term is still being determined. Further close monitoring of these schemes is needed so that adjustments can be made where necessary to improve service delivery and strengthen the prospects for sustainability. Lessons arising from this are likely to be of interest to the sector.**

ES4.4 Recommendations for Independent Process Monitoring and Evaluation

While the evaluation did not extend to the IPME team evaluating itself, reflection on the practice of independent process monitoring and evaluation under ASWA suggests lessons for a future IPME (or its equivalent).

1. IPME should be in place at the start of the ASWA programme to support and quality assure baseline surveys, and conduct monitoring systems appraisals much earlier. Subsequent output verification should occur at agreed intervals. In addition, to secure timely action on recommendations, six-monthly tripartite meetings between UNICEF, DFID and IPME should review latest IPME reports and agree on any remedial actions.

2. IPME should maintain a change log of agreed recommendations and remedial actions, with progress to be updated before six-monthly tripartite meetings to assist programme oversight.

3. IPME's programme director and UNICEF HQ's ASWA programme manager should hold virtual discussions between six-monthly tripartite meetings to conduct a health check on the IPME – UNICEF relationship and consider any recommendations for improvement to be discussed with DFID.

1. INTRODUCTION

1.1 Introduction

WYG International (WYG) in association with Aguaconsult is pleased to submit this final evaluation report of the Accelerating Sanitation and Water for All (ASWA) Programme in Neglected, Off-Track Countries. The evaluation has been conducted as part of the Independent Process Monitoring and Evaluation (IPME) contract from the UK Department for International Development (DFID) held by WYG[2]. This report presents findings and conclusions gained through implementation of the evaluation design as set out in the IPME ASWA Programme Evaluation Design Document (see Annex S). This report is presented to DFID – as funder of ASWA – and to UNICEF.

1.2 Purpose of the ASWA Final Evaluation

The main purpose of the evaluation as originally set out in the IPME Terms of Reference (TOR) (DFID, 2013b – see also Annex M) was that "...the impact evaluation element will address key questions for the design and implementation of this and future WASH sector programmes". During the inception phase, it was agreed with DFID that the evaluation purpose would be revised as follows:

> "To investigate the reasons behind the achievement/non-achievement of verified results of the ASWA program and to gain a deeper understanding on the functioning of selected ASWA country programmes" (IPME 2015, p. 31)

The IPME Final Inception Report (IPME, 2015) also set out that the evaluation would not aim to measure ASWA's impacts (as suggested in the IPME TORs). Rather, it would assess both the quality of the outputs (facilities and behaviours) and the prospects for sustainability, which is a pre-requisite for the impacts sought by ASWA (see IPME 2015, p. 4). Furthermore, due to budget constraints it was agreed during the inception phase that it would not be possible to evaluate all the nine countries of the ASWA programme to the same depth (IPME 2015, p. 31). The original provision for evaluation activities in four countries has, as foreshadowed in the Inception Report, been revisited and set at activities in three countries.

Given the revised purpose and country coverage, the focus of the evaluation is on learning within and for ASWA country programmes, rather than more widely across all of UNICEF's and DFID's programmes in the WASH sector. The evaluation's main concern is not to evaluate a generic Theory of Change (TOC) for WASH interventions. It does aim to understand why some country programmes have performed better than others, and within country programmes to explore why some implementing partners have performed better than others. In doing so, it seeks to understand how specific local contexts have affected programme outcomes.

In terms of intervention design, the nine ASWA country programmes are not seeking to break new ground, rather they are applying tested operational approaches – not least Community-Led Total Sanitation (CLTS) which is one of the most commonly used methods for rural sanitation and hygiene promotion in developing countries. For DFID, the most important question is how effectively UNICEF performed in delivering the ASWA programme using CLTS and other familiar WASH interventions. Hence the success or failure of programme implementation (which has a lot to do with programme management) receives primary attention; we are only secondarily focusing on the intervention theory. Given this situation, issues of external validity in relation to the findings from the evaluation are not of paramount importance in this type of evaluation. Where the evaluation analysis corroborates or

[2] For further information on the IPME contract see Annex M: IPME TOR.

challenges findings from other studies in the WASH sector, the report notes this. But we have been careful not to make claims of wider relevance where these cannot be robustly supported.
Three other significant areas of focus requested by DFID in the evaluation were: focusing on issues of outcomes, sustainability, and value for money; assessing prospects for sustainability rather than impact directly; and exploring UNICEF mechanisms, systems and processes for gathering gender and other disaggregated data[3]. In short, the objectives of the ASWA Final Evaluation, in order of priority for DFID, are to understand and share lessons on:

- The prospects for sustainability of outcomes.
- The extent of value for money.
- How and why verified results were achieved / not achieved.
- The quality of outputs.

It is important here to clarify how evidence for outcomes differs from that for outputs. While programme results at output level are primarily concerned with 'access' to WASH facilities (an exception being the ODF indicator), outcome indicators are concerned with their 'use'. Specifically, the use of sustainable sanitation facilities, the use of sustainable water services and the adoption of hand washing with soap and water after defecation. Outputs can generally be tracked through routine monitoring and reporting by Implementing Partners (IPs), but the assessment of programme outcomes requires special studies to assess changes in personal behaviour at endline compared to baseline. **Virtually no relevant outcome data was available by 30 June 2016[4] for the majority of ASWA programme countries** and this inevitably affected the scope of the evaluation analysis and findings. (See Section 3.4.1 below for a full discussion of how the evaluation sought outcome evidence).

1.3 Scope of the Evaluation

The subject of the evaluation is the ASWA Programme at country level (including support from UNICEF Head Quarters [HQ] and Regional Offices [ROs]) between November 2013 and September 2016. The ASWA Programme was extended beyond September 2016 during the evaluation, but it was agreed with DFID that the evaluation would not cover this extension period. The cut-off point for data considered by the evaluation was set at June 2016 (See Section 1.5 below for rationale).

The scope of the evaluation in relation to Outputs 1-4 is the contribution of the results of the DFID-funded ASWA programme specifically. However, for Output 5, which deals with the WASH sector enabling environment, the focus is broader to include the contribution of all UNICEF WASH activities at the country level, including support from UNICEF HQ and ROs[5]. Furthermore, because ASWA is designed with a multi-actor sector approach, evaluation of attribution was not possible or desirable (as this would artificially privilege ASWA when UNICEF support to government systems using several funding sources and coordination with other development partners was integral to UNICEF's approach)[6].

The original evaluation tasks envisaged within the wider TORs for IPME (see Annex M) are compared to what happened in practice, with justifications for any departures in Table 1 below.

[3] For further information on why impact evaluation was not sought see Evaluation Design Document (Annex S), p. 49.

[4] The cut-off date for evidence to be considered by the evaluation (see Sec 1.5 below).

[5] The ASWA Outputs defined in the DFID ASWA Revised Logframe (2015) are: Output 1: 5 million people live in Open Defecation Free Communities, of which 1.5 million gain access to sustainable basic sanitation facilities; Output 2: 0.9 million people will gain access to sustainable improved water supplies; Output 3: 10 million will be reached by hygiene education programmes of which 4.5 million people adopt habitual Hand Washing with Water and Soap (HWWS); Output 4: 400,000 school children will benefit from improved WASH facilities in schools; and Output 5: WASH sector enabling environments strengthened through the removal of key institutional barriers to progress; capacity building for lead sector institutions; and enhanced learning on operational knowledge gaps.

[6] For further information on contribution vs. attribution see Evaluation Design Document (Annex S), p. 47 and p. 51.

Table 1: Original Evaluation Tasks Compared to Implementation

Tasks from IPME TOR	Evaluation Implementation	Justification for Any Departures
Developing an inception and work plan for process and impact evaluation (including literature review on WASH effectiveness, evaluation design and collection of data, risk management, quality assurance, evaluation questions, strategies for linking with other initiatives and dissemination).	Inception Report produced (IPME 2015, Inception Report), and full Evaluation Design separated out and produced later (see Annex S, Evaluation Design).	Evaluation section of Inception Report was high level and required further consultative development and detail.
Designing and implementing process evaluations to inform assessment of programme performance and to highlight areas for improvement during the programme.	Assessment of ASWA Annual Reviews in 2015 and 2016.	Synergy with the IPME Monitoring and Verification work stream feeding into the Global Review of Programme Results method.
Designing and implementing impact evaluations to address innovative elements of the programme and evidence gaps in relation to sustainable provision of WASH services.	Impact evaluation was not attempted, but rather the evaluation assessed both the quality of the outputs (facilities and behaviours) and the prospects for sustainability, which is a pre-requisite for the impacts sought by ASWA.	Over the course of the inception period both the IPME team and DFID jointly agreed that this requirement is unlikely to be possible for three main reasons: 1) The project has a short time frame and is due to end just three months after the delivery of outputs, which is too soon to expect any measurable impact; 2) To measure impacts rigorously, specially commissioned baseline studies would have been needed at programme inception (before IPME involvement) but no provision was made for this; 3) In the case of health, it is notoriously difficult to measure the impact of WASH interventions. We note here that substantial resources were deployed to measure the health impact of the much longer and larger SHEWA-B programme in Bangladesh, but none could be detected[7].
Tracking whether assumptions set out in the Theory of Change and log frame hold (including expanding on the Theory of Change in relation to relevant programme issues such as enabling environment, governance and sustainability).	An elaborated ASWA Theory of Change, developed by IPME in consultation with DFID and UNICEF, informed the evaluation questions, Theory of Change Review method, and analysis of programme logic.	No departure.
Evaluating innovative elements of the programme.	Evaluation questions on innovation were included under the evaluation criteria Upscaling.	No departure.
Developing the evidence base on value for money metrics, comparing suppliers and approaches across contexts.	A Programmatic Value for Money Assessment was conducted as part of the final evaluation.	No departure.

[7] Sargsyan et al (2014), Bangladesh WASH sector: large scale impact assessment, Briefing Paper 1983, WEDC, Loughborough University, Loughborough.

Tasks from IPME TOR	Evaluation Implementation	Justification for Any Departures
Identifying factors which have enhanced or impeded the sustainability of WASH interventions.	Enabling and constraining factors to sustainability were included within the evaluation questions.	No departure.
Addressing knowledge gaps (e.g. effective approaches to behaviour change, community monitoring and payment by results).	Knowledge gaps in the WASH sector were identified in the evaluation design and informed the evaluation questions.	No departure.
Assessing how the wider environment has enabled or impeded achievement of programme objectives and identifying implications for programming.	Evaluation questions exploring potential rival explanations for observed ASWA outcomes were included in the evaluation questions.	No departure.
Identifying regional partners with potential to provide evaluation services at national, district and community level.	IPME undertook a scoping study to identify potential regional partners to support sector monitoring and strengthen the accountability of government and development agencies to local stakeholders in ASWA countries. This aspect of the TORs was not taken further.	Variation of TORs agreed with DFID as part of the wider IPME contract.
Considering instruments already in use and others which can be adapted at country level to build and strengthen indigenous evaluation procedures drawing on a selection of international examples.	See above.	See above.
Dissemination of evaluation findings potentially including a peer review publication, presentations and international meetings e.g. Sacosan / Africasan, Sanitation and Water for All High-Level Meeting, Stockholm World Water Week.	A global learning workshop to support validation of the evaluation findings and to share wider learning from the IPME contract is envisaged, but needs assessment and agreement with DFID and UNICEF is ongoing.	No departure.

1.4 Evaluation Questions

Twelve priority evaluation questions, and five secondary evaluation questions were developed by IPME in a collaborative process involving the primary stakeholders (DFID and UNICEF) through a series of iterations. The final evaluation questions are informed by the revised DFID ASWA logframe targets (see Annex C), elaborated ASWA TOC (see Annex B), OECD-DAC evaluation criteria, UNICEF evaluation criteria (see Annex N), knowledge gaps in the WASH sector, and evidence already flowing from other IPME work streams (Monitoring and Verification; Customised Support) identifying important assumptions and contextual factors.

The original ASWA TOC, as set out in the ASWA Business Case (DFID, 2013a) provided an important starting point for the development of evaluation questions, but had several limitations[8] as a programme theory for ASWA. The evaluation team therefore engaged DFID and UNICEF through an iterative process as part of the evaluation design and early implementation to define an elaborated ASWA TOC. DFID indicated that, for the purposes of the evaluation, the TOC should focus on linking mechanisms and assumptions at input, output and outcome stages (it therefore excludes linking mechanisms and assumptions between outcome and impact stages). The TOC also focuses on the effectiveness of

[8] These are set out in the Evaluation Design (see Annex S) but in short, the original TOC: did not fully represent the evidence set out in the Business Case for linking mechanisms that explained transitions between inputs, outputs, outcomes and impact; did not make explicit where the assumptions in the Business Case fit within these relationships; and was not fully aligned with the ASWA logframe, which was developed after the TOC.

UNICEF in achieving sustainable gains and working across the sector to strengthen enabling environments and systems for ongoing service delivery/behaviour change. As such this is not a comprehensive 'programme theory' evaluation as it does not attempt to evaluate all aspects of the TOC. The elaborated ASWA TOC agreed with DFID and UNICEF is set out in detail in Annex B and highlights the assumptions and linking mechanisms that the evaluation did and did not seek to explore in line with DFID's areas of focal interest.

While the IPME TOR suggested that evaluation questions should be in line with OECD-DAC evaluation criteria[9], consultations with UNICEF suggested the questions should also align with criteria developed by UNICEF's Evaluation Office for WASH programmes which build on those of the OECD-DAC (see Annex N). The combined OECD-DAC / UNICEF evaluation criteria used are Relevance, Effectiveness, Efficiency, Sustainability, Equity and Up-scaling. The impact criterion does not appear in the list because sustainability is being used as its proxy (see Section 1.1 above).

The ASWA Evaluation Questions are shown in Table 2.

Table 2: ASWA Evaluation Questions[10]

No.	Primary Evaluation Question	Secondary Evaluation Question
Combined OECD DAC / UNICEF Criteria: RELEVANCE		
1	Is the design of each ASWA country project relevant to and coordinated with the on-going or planned strategy, impacts, outcomes and outputs being provided by national government and other key development partners (including NGOs); and with DFID's and UNICEF's country strategy / plan?	1.b How is UNICEF HQ (Programme Division WASH) introducing guidance and technical assistance on climate change adaptation and WASH to ASWA country offices?
	1.a Were the needs of target populations identified and used to inform the design of ASWA at country level?	
Combined OECD DAC / UNICEF Criteria: EFFECTIVENESS		
2	To what extent and why are people (in communities and schools): • Using and maintaining sanitation facilities? • Adopting hand washing with soap / ash and water after defecation? • Using and maintaining water services?	2.b To what extent have outputs been delivered and outcomes likely to be achieved?
	2.a What were the key enabling and constraining factors to delivery of outputs and likely achievement of outcomes at the level of UNICEF (e.g. partnership arrangements and management, procurement strategy), implementing partners and communities / households?	
3	To what extent are donors and other development partners willing to coordinate and engage in joint learning to improve the WASH sector at national level?	3.b Has ASWA leveraged additional public financing to sustainably support strategic gaps in the WASH sector at all levels?
	3.a How effectively is UNICEF engaging with government to strengthen WASH sector coordination at national level?	
4	To what extent has ASWA improved enabling environment for WASH at national and / or sub-national levels and what has been its approach (e.g. in terms of sector systems and monitoring, strengthening government led scale up efforts, local government promotion / incentives and regulation, evidence based policy, and clarifying roles and responsibilities at decentralised levels)?	
	4.a How is government leadership and political will influenced to improve the WASH sector?	
	4.b To what extent is there local and national government capacity in key areas of sustainable WASH implementation?	
Combined OECD DAC / UNICEF Criteria: EFFICIENCY		

[9] The OECD Development Assistance Committee or DAC criteria for evaluating development assistance programmes are well recognised and established as a framework for carrying out the type of evaluation as set out in the IPME contract; see: http://www.oecd.org/development/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm.

[10] For further information on the relationship between the EQs and the TOC, and the relationship of the EQ with knowledge gaps, please see (see Annex S, Evaluation Design Document, pp. 71 – 80).

No.	Primary Evaluation Question	Secondary Evaluation Question
5	Is the programme delivering Value for Money in terms of effectiveness, and equity and why?	5.b Is the programme delivering Value for Money in terms economy and efficiency and why?
	5.a What are the evidence streams to support the conclusions reached about Value for Money?	
Combined OECD DAC / UNICEF Criteria: SUSTAINABILITY		
6	To what extent are the benefits of the ASWA programme likely to continue after DFID funding ceases?	
	6.a How appropriate are the locations for water supply services?	
	6.b How willing are communities to participate in and mobilise for ODF?	
	6.c How willing are communities to sustain ODF?	
	6.d What is the level of quality of construction of water supply and sanitation services (in communities and schools)?	
	6.e What is the level of quality of supervision of operational water supply services (in communities and schools) by community level WASH committees and authorities supervising water service providers?	
7	What are the major factors and drivers influencing the likely achievement or non-achievement of sustainability of the ASWA objectives for · Sanitation Facilities? · Water services? · Hygiene?	
	7.a To what extent is finance in place to sustainably support WASH sector systems development beyond external programme funding?	
	7.b How adequate is government support to decentralised administrative levels?	
	7.c To what extent is the local private sector able to function profitably and willing to engage (e.g. in supplying services to hygiene promotion and / or operation and management of water supply)?	
	7.d To what extent are investments in WASH sector learning, coordination and other intangible activities valued by government and development partners?	
	7.e How effectively is UNICEF embedding community-led monitoring and the role of civil society in monitoring and accountability systems?	
Combined OECD DAC / UNICEF Criteria: EQUITY		
8	How appropriate was the selection of communities and schools in terms of targeting benefits at communities in the lowest wealth quintile and at women and girls?	8.b To what extent are processes and tools in place to identify, target and monitor equity between population groups with different social characteristics based on their needs?
	8.a To what extent has UNICEF's approach to promoting equity been operationalised and successful or not?	
Combined OECD DAC / UNICEF Criteria: UP-SCALING		
9	What other changes (positive / negative, direct / indirect; intended/unintended) have occurred as a result of ASWA interventions?	
	9.a What evidence is there that particular innovations within the ASWA programme are being replicated beyond the initially intended reach of the programme (e.g. outside of geographic areas or target groups)?	
Context		
10	Are WASH programmes and / or humanitarian programmes with WASH components being delivered by other development partners in the same locations as ASWA with related outputs and outcomes?	
11	Have natural disasters or conflict prevented or impaired project implementation or damaged water or sanitation facilities post construction?	
12	Have natural, political, economic forces or pandemic disease led to the displacement of target populations?	

Further iteration of the evaluation questions was undertaken on an ongoing basis during the evaluation's implementation by verifying and clarifying important assumptions and contextual factors. Prior to each case study country evaluation mission, country-specific interview questions based on the evaluation questions to be addressed by the Key Informant Interview (KII) method were developed. These considered emerging findings from sustainability assessments in each country, IPME monitoring and verification reports, advice from IPME country leads, and contextual information provided by the evaluation team national consultants. A draft of country-specific interview questions was verified by DFID and the relevant UNICEF Country Office prior to finalisation. During each evaluation mission the implementation of the country-specific interview questions was focused to pursue emerging findings and contextual issues identified from the Rapid Outcome Assessment (ROA) and Theory of Change (TOC) workshops in each case study country.

1.5 Evaluation Timing

The evaluation covers the period from the start of the programme in November 2013 until September 2016[11]. The evaluation design (see Annex S) was agreed with DFID in May 2016, in preparation for the in-country evaluation visits to Myanmar, Pakistan and Madagascar in June, August, and September 2016, respectively. At the time of planning for the implementation of the evaluation (Quarter 1 2016), the ASWA programme was due to end in March 2017. Staging of the evaluation was therefore timed so that it would report as late as possible into programme implementation, when most of the outputs had been delivered.

In the latter half of 2016, however, both ASWA and IPME were extended (to March 2018 and November 2017, respectively). Nevertheless, the timing of the evaluation and cut-off point for data of June 2016 remained unchanged for three reasons:

- ASWA programme level results to June 2016 for all outputs had already exceeded the Milestone 2 (2015) levels and it was not justifiable to wait for more results that would further exceed these levels.
- By June 2016 outcome data was unavailable for most ASWA countries. Other studies planned by UNICEF were either not ASWA-specific and / or would not have access to baseline data that would be a pre-requisite to assess progress on outcomes at a later stage.
- Significant elements of the evaluation missions had already been implemented.

1.6 Evaluation Stakeholders

The primary audience for the evaluation (as envisaged by the TOR [DFID, 2013b: p. 4]) is DFID and UNICEF. The evaluation design (see Annex S) further specified the main stakeholders in the evaluation as being: DFID Water, Sanitation and Hygiene Team and DFID Country Offices; UNICEF Headquarters WASH Team, Headquarters Evaluation Office, Regional Offices and Country Offices. Other interested development partners include Government Counterparts, Implementing Partners, Service Providers, Banks / Chambers of Commerce, Academics, Journalists, Rights-Based Organisation, Faith-Based Organisations, Local Councillors, Community Based Organisations, Communities and Households[12].

1.7 Structure of this Report

The evaluation report has been designed to keep the main body as short and readable as possible for the primary audience (DFID and UNICEF). This necessarily means that most of the evidence collected

[11] Given that the evaluation missions were ongoing between June – September 2016, country and programme data provided by UNICEF is to the end of June 2016. COs report six monthly, so data for the period July to September 2016 was not due to be produced before early 2017. In the event, the December 2016 reports were not shared with DFID and IPME until July 2017.

[12] For further information on the envisaged roles of stakeholders please see Evaluation Design Document (Annex S), pp. 52-53.

and analysis carried out is presented in supporting Annexes. Those seeking such detail should consult the relevant Annexes (which are referenced in the main body of the report where appropriate). The findings and conclusions (Section 3) are presented by ASWA output or theme instead of by evaluation question to enhance the utility of the evaluation report. Those seeking a presentation by evaluation question should consult Annex A. **Recommendations for DFID, UNICEF, the WASH Sector and IPME are presented in the Executive Summary above (see Section ES4).**
The remainder of this Final Evaluation report is structured as follows:

Section 2: ASWA Programme Design and Implementation – Implementation of the ASWA programme in terms of its context, design and content, and results to date.

Section 3: Findings and Conclusions – Evaluation findings and conclusions by ASWA output or topic in terms of results achievement, quality of outputs, outcomes and sustainability, value for money, equity, programme monitoring, and other changes and innovations.

Section 4: Evaluation Approach – Evaluation design, process of implementation, methods and revisions, limitations, and issues of inclusion and research ethics.

Section 5: References – Documents referred to in the main body of the report.

Appendix 1: Programme Logic Tables - Findings and conclusions on the programme logic from related portions of the elaborated ASWA TOC

Annexes
All Annexes are publicly available online and can be accessed from the links given for each Annex below.

A – Table of Findings and Conclusions by Evaluation Question (see for an alternative summary presentation of material from Section 3). Annex A[13]

B – ASWA Theory of Change (see for the overall programme logic explored through the evaluation). Annex B[14]

C – ASWA Logframe (see for the original and revised DFID description of ASWA impact, outcomes, outputs, and targets). Annex C[15]

D – Evaluation Framework (see for evaluation questions, information needed / source of data, methods of analysis, and standard for judging performance). Annex D[16]

E – Madagascar Country Case Study (see for full report on embedded Madagascar case study). Annex E[17]

F – Myanmar Country Case Study (see for full report on embedded Myanmar case study). Annex F[18]

[13] https://www.dropbox.com/s/r67z34ivm1f7aet/AnnexAFindingsandConclusionsbyEQ.pdf
[14] https://www.dropbox.com/s/6p45k2gd345hafr/AnnexBASWATheoryofChange.pdf
[15] https://www.dropbox.com/s/chmwowbrjt0wuf7/AnnexCASWALogframe.pdf
[16] https://www.dropbox.com/s/czwuobv3dpoqosf/AnnexDEvaluationFramework.pdf
[17] https://www.dropbox.com/s/z59qhzj18apyy1a/AnnexEMadaEvnCS.pdf?dl=0
[18] https://www.dropbox.com/s/x05qju54ajkusjs/AnnexFMyanEvnCS.pdf?dl=0

G – Pakistan Country Case Study (see for full report on embedded Pakistan case study). Annex G[19]

H – Cross-Country Case Study Analysis (see for triangulation of evidence from embedded country case studies for analysis across Madagascar, Myanmar and Pakistan ASWA country programmes). Annex H[20]

I – Global Analysis (see for triangulation of evidence from global level methods for analysis across all nine ASWA country programmes). Annex I[21]

J – Analytical Tables (see for triangulation by method for: Global Review of Programme Results (providing evidence for Annex I Global Analysis); Madagascar, Myanmar, and Pakistan Country Case Studies (providing evidence for Annex H Cross-Country Case Study Analysis). Annex J[22]

K – Documents Reviewed (see for full list of Documents Reviewed by the Evaluation). Annex K[23]

L – Operational Guidance Notes (see for protocols that guided the implementation of evaluation methods). Annex L[24]

M – IPME TOR (see for the original DFID Terms of Reference for the ASWA Independent Process Monitoring and Evaluation team that encompass this final evaluation). Annex M[25]

N – UNICEF Evaluation Office Evaluation Criteria and Key Questions for WASH (see for evaluation criteria complimenting OECD-DAC that were combined to inform the evaluation design and questions). Annex N[26]

O – Sustainability Assessment Myanmar (see for full report on the sustainability assessment carried out by this Final Evaluation as part of the Myanmar country case study). Annex O[27]

P – Programmatic VFM Study (see for full report on Programmatic Value for Money Assessment of ASWA carried out by this Final Evaluation as part of Global Review of Programme Results). Annex P[28]

Q – List of Consultees (for organisations and groups consulted by this Final Evaluation). Annex Q[29]

R – ASWA Evaluation KII Method Level Word Table (see for example of analytical table used at method level to collate and analyse evidence). Annex R[30]

S – Evaluation Design (see for detailed technical design of the evaluation). Annex S[31]

[19] https://www.dropbox.com/s/jte5sfl4q42atih/AnnexGPakEvnCS.pdf?dl=0
[20] https://www.dropbox.com/s/24gdzqhx04d5on0/AnnexHCrossCountryCaseStudyAnalysis.pdf
[21] https://www.dropbox.com/s/1t16qrrn24b6tih/AnnexIGlobalAnalysis.pdf
[22] https://www.dropbox.com/s/ex59y13wqrqer7d/AnnexJAnalyticalTables.pdf
[23] https://www.dropbox.com/s/fs7c5f1htlci277/AnnexKDocumentsReviewed.pdf
[24] https://www.dropbox.com/s/ij2dnqjf6tjusnm/AnnexLOperationalGuidanceNotes.pdf
[25] https://www.dropbox.com/s/arbmicty1tgkvhz/AnnexMIPMEToR.pdf
[26] https://www.dropbox.com/s/qe0la3lbmedh5as/AnnexNUNICEFEvnOfficeEvnCriteria%26KeyQsforWASH.pdf
[27] https://www.dropbox.com/s/nwkm3laz2t0q1zy/AnnexOSAM.pdf
[28] https://www.dropbox.com/s/jzd78mfzp26bm34/AnnexPProgrammaticVFMStudy.pdf
[29] https://www.dropbox.com/s/k37brf8aasxanh8/AnnexQListofConsultees.pdf
[30] https://www.dropbox.com/s/p4g50qgc31dj99y/AnnexRKIIMethodLevelWordTable.pdf
[31] https://www.dropbox.com/s/toif1452egm8sf6/AnnexSEvaluationDesign.pdf

2. ASWA PROGRAMME DESIGN AND IMPLEMENTATION

2.1 Introduction

This section describes the ASWA Programme's design and implementation by DFID and UNICEF to provide readers unfamiliar with ASWA a background understanding of the programme's context. It also provides an overview of how the IPME contract relates to the programme, how the programme was designed, the programme's content (including logframe, country level and cumulative programme level targets and results, budget distribution) and results to date.

2.2 Context

According to the DFID Business Case (DFID 2013a), ASWA has its origins in a UK government commitment made at the Sanitation and Water for All High-Level Meeting in April 2012, at which the government committed to double the number of people it would help gain access to water, sanitation and hygiene promotion to 60 million people by 2015 (the Millennium Development Goal [MDG] target date). DFID already provided WASH support to many countries, much of it through bilateral agreements, but it was estimated that these programmes would be unable to deliver the 60m beneficiaries target on their own. ASWA was formulated as an additional contribution to help plug the gap through a grant of £35.5 million to UNICEF.[32]

The stated intention of DFID support was to accelerate progress towards meeting MDG target 7c (to halve by 2015 the proportion of the world's population without sustainable access to safe drinking water and basic sanitation) in those countries where progress had been slow, and to complement related MDGs in boosting child survival. ASWA would result in improved access for at least 5 million people to sustainable WASH services in off-track countries. UNICEF would propose up to 12 countries where the programme would be implemented, based on selection criteria including: numbers of people lacking access, country income level, state fragility, extent to which the country was off-track to meet the MDG targets, and level of inequity in access to water and sanitation.

The ASWA Business Case highlighted that progress towards the MDG target for sanitation had been slow, leaving it one of the most off-track of all the targets. In contrast, the MDG water target had been met at the global level five years ahead of schedule. However, this progress masked considerable disparities in access within and between countries. Sub-Saharan Africa was highlighted as the region most off-track for both drinking-water and sanitation, while South Asia was off-track for sanitation and was the region with the highest number of people without access to any kind of sanitation facilities. The ASWA programme was to focus on off-track countries in these regions. The Business Case also said that "the UK would focus on the barriers that hold back universal WASH access":

- ensuring that water and sanitation infrastructure is well maintained and sustainable;
- reaching the poorest in difficult to reach areas and those who have limited ability to pay;
- developing the evidence base on effective models for service delivery at scale;
- mobilising private sector investment and expertise; and
- building demand for sanitation and hygiene amongst the poor and facilitating access."

2.3 IPME Engagement

Alongside the grant to UNICEF, DFID contracted through a competitive bidding process involving UNICEF, a third-party organisation (WYG International) for Independent Process Monitoring and Evaluation (IPME) of the ASWA programme. Key tasks for IPME included "quality assuring programme

[32] The WASH Results Programme, which ran concurrently with ASWA, was also conceived to help DFID deliver on its 60 million target.

progress reports, ensuring that robust baseline data is collected by the programme and assessing whether results reported are attributable to the programme." The scope of work also included the provision of demand-responsive technical support to strengthen UNICEF's programme monitoring capacity and to support UNICEF COs' work with government to strengthen sector monitoring, as well as this final evaluation (see Annex M IPME TOR).

The IPME team's prior engagement with the programme had two key implications for the evaluation:

1. Programme results in six countries had been validated ahead of the evaluation missions.
2. Members of the IPME team had gained insights on programme operations over the preceding two years, so had some knowledge of how programme results had been delivered.

The evaluation did not, therefore, start with a blank sheet and has drawn on findings from this earlier engagement. Nevertheless, there has been more engagement with some countries than others, as the extent of support requests varied. Security restrictions also prevented the IPME team from making field visits outside the capital in South Sudan and Niger, while Yemen could not be visited at all.

2.4 Programme Design and Content

The detailed design of the ASWA Programme took place over an extended period of discussion and negotiation between DFID and UNICEF HQ up to its launch in August 2013. UNICEF's preference at the outset had been for thematic funding, meaning that the resources would be provided to support UNICEF's global WASH programme in general without being tied to specific outputs. DFID did not agree to this, however, since it would be hard to ensure that the funds were spent on DFID priorities. In addition, DFID had concerns with the quality of UNICEF's internal monitoring and reporting processes; tying funding to an implementation project offered more leverage to ensure that these systems were improved[33].

DFID's broad intention was for ASWA to operate in countries where it had no existing bilateral WASH programme, though Bangladesh was in fact allowed as an exception. The final selection of countries is presented in Table 3 below[34]:

Table 3: ASWA Regions and Countries

Region	Countries
Sub-Saharan Africa	Madagascar, South Sudan, Niger
South Asia	Pakistan, Nepal, Bangladesh
South-East Asia	Cambodia, Myanmar
Middle East	Yemen

Two of the countries (Niger and Madagascar) were not receiving any DFID support when the programme was planned. Amongst the others, some were in receipt of DFID humanitarian support in WASH while DFID's WASH Results Programme (which, like ASWA, is centrally managed by DFID) would later operate in four: Pakistan, Bangladesh, Nepal and South Sudan. Seven of the countries were included in a table of 29 countries currently supported by DFID which appeared in the Business Case. All were listed as being off-track for sanitation, while only two (South Sudan and Yemen) were shown as off-track for water supply.

In each country, ASWA was a sub-set of a wider UNICEF country programme designed in close collaboration with national government – the typical approach for UNICEF country programmes globally. Although the evaluation was not tasked with assessing the ASWA programme in relation to the Paris

[33] Interview with DFID WASH Policy Team, January 2017.

[34] For further information on the selection of ASWA countries see Annex P Section 2.6, p. 13.

Declaration[35] principles for aid effectiveness, the following observations are made. The ASWA programme reflects four of the five Paris Declaration principles: country ownership (e.g. UNICEF country programmes are jointly owned by national government); donor alignment behind country objectives and systems (e.g. ASWA sought to support and / or strengthen national WASH sector policies and systems); harmonisation between DFID and UNICEF (e.g. ASWA was grant funding that supported WASH in existing UNICEF country programmes rather than DFID bilateral spend in parallel to UNICEF); and a focus on results and their measurement (e.g. a results-oriented logframe supported by independent third-party process monitoring and evaluation). The principle of mutual accountability between donors and partner countries is less evident (e.g. UNICEF reports ASWA performance to DFID, but governments in ASWA countries are not directly involved in this accountability process).

The original ASWA logframe is provided in Annex C. UNICEF (primarily HQ) was heavily involved in developing the logframe with DFID, and both organisations agreed to the targets it contained. The logframe described a broadly standard WASH programme with hygiene promotion and WASH in schools components in addition to household sanitation and community water supply. In this respect ASWA represented 'business as usual' for UNICEF. The programme was not introduced to break new ground in terms of objectives or the means of achieving them, though it was expected to deliver results (defined as numbers of beneficiaries) on a fairly large scale overall and within a fairly short timeframe.

It is important to note that only a global logframe was developed, with no disaggregation of targets by country. Country-specific components and targets were negotiated between UNICEF HQ and COs. It is useful to highlight some other features of the original logframe that are salient to the evaluation:

1. With only global targets listed, the implication was that shortfalls in results from one country could be compensated for by over-delivery in another.

2. The logframe appeared to anticipate an equal number of people gaining access to sanitation and living in ODF communities[36]. It is very unlikely in practice that these numbers would be equal within individual country projects, as there is often some level of latrine coverage at baseline and not all targeted communities become ODF.

3. The logframe anticipated 100% conversion of outputs into outcomes (i.e. transition from access to use of facilities and from receipt of hygiene promotion to the adoption of hand washing with soap and water after defecation). Sector experience indicates that this ambition was unrealistic. We note here that in DFID's WASH Results Programme, country-specific outcome level targets were defined ahead of the second implementation phase, though not at the programme start. For sanitation, 70% to 75% of declared ODF populations continuing to use toilets was considered acceptable, as was an improvement of 10-15% over baseline in people practising hand washing with soap. Under ASWA, COs were not required by DFID or UNICEF headquarters to define outcome level targets in their funding proposals, nor was this done later.

4. There was no requirement for the programme to achieve universal access across entire sub-districts or districts; results were defined only in terms of the total numbers of beneficiaries.

[35] The Paris Declaration sets out five principles that were endorsed by governments at the second High Level Forum on Aid Effectiveness (2005). For further information see: www.oecd.org/dac/effectiveness/parisdeclarationandaccraagendaforaction.htm.

[36] Output 2 read: "5 million people gain access to sustainable basic sanitation and live in ODF communities," though the associated indicators gave different targets for ODF and for gaining access.

5. Since the logframe defined only global targets and included no baseline data, it was impossible to discern the level of ambition of the programme at country level in terms of improvement over baseline.
6. Under the water supply output, there was an indicator for 'water resource and climate change assessments' undertaken at watershed level prior to the construction of new supplies, with a target to complete 864 by programme end. What constituted such an assessment, however, was not made clear in the logframe or Business Case. At the planning stage, UNICEF had no methodology for these assessments and DFID accepted that the design of each one would be context-specific. In practice, this indicator proved problematic as COs were unsure what they were expected to do (see Section 3.2.7.1).

7. The logframe included (at UNICEF's request) target numbers of indirect water supply and sanitation beneficiaries. The Business Case indicated that these would be delivered through 'system strengthening,' meaning interventions to enhance the enabling policy and institutional environment.

8. Output 5 was concerned with strengthening the enabling environment and contained five indicators, but programme ambitions in this area were not well defined. In particular:

    - The logframe showed the baseline as zero for each indicator, when this was not in fact the case. For example, against indicator 5.1, several of the ASWA countries (for example Nepal, Madagascar) already had sector policies and strategies and / or had introduced a system of joint annual or bi-annual sector reviews with mechanisms for corrective action. This anomaly arose because the logframe was written before the countries had been selected. The logframe anticipated that only some of the nine countries would deliver results against each indicator, but UNICEF headquarters did not facilitate an agreement at programme start as to which countries would pursue which indicators. In the event, while all country programmes had a substantive enabling environment component, ASWA-specific objectives remained undefined in most cases (Bangladesh being the exception).

    - Indicators 5.1 to 5.4 described long term ambitions that were unlikely to be delivered by UNICEF support alone, or within the time available. As such they were not very useful as indicators of progress within the ASWA timeframe.

    - Indicator 5.5 (institutional capacity strengthened) was framed imprecisely, so it would be difficult to assess whether it had been achieved.

    - Under indicator 5.6, UNICEF HQ planned to use a portion of programme funding to undertake several WASH-related research projects in non-ASWA countries. This was a stand-alone initiative largely disconnected from the rest of the programme. It was not clear to what extent the ASWA country programmes were also to undertake research.

When the IPME assignment began in May 2014, it became clear that the above points made assessing progress against the logframe problematic. IPME inception visits to seven countries from late 2014 to early 2015 found a lack of common understanding of the logframe content within UNICEF. It appeared that there had been insufficient dialogue between headquarters and COs on the logframe. Some key staff had not even seen the final version of the logframe, though had seen earlier drafts. Had country-specific logframes been formulated by UNICEF, some of these challenges might have been avoided.

In response, the IPME team worked with DFID to improve the logframe in consultation with UNICEF, and a revised version was adopted in January 2015 (see Annex C). The revisions were an attempt to

give the outputs more focus and make the indicators more verifiable. In addition, a completely new indicator was added, for piloting a methodology for the measurement of indirect beneficiaries. This replaced outputs 2.1 and 3.2 (beneficiaries delivered through 'system strengthening') which had proved unworkable since UNICEF had not (so far) developed a viable means of counting indirect results.

2.4.1 Country Programme Focus and Targets

Irrespective of these modifications, COs have approached the logframe on the basis that, since programme targets were skewed towards sanitation and hygiene promotion, they should prioritise Outputs 1 and 3. The other outputs have been treated as a menu of options and the extent to which COs pursued them depended on their perceived local relevance and/or the availability of alternative funding sources. In the case of output 5, for example, some COs had ongoing initiatives related to one or more of the indicators which began prior to ASWA and were funded from other sources, so this activity was not always mentioned in ASWA reports.

The scale of the targets adopted under Outputs 1 to 4 varied considerably across the nine countries[37]. For sanitation, targets ranged from just 42,000 in South Sudan and 57,000 in Yemen to 970,000 in Madagascar and 1.2 million in Nepal. Cambodia did not include a water supply component in their programme and for the others, the range of water supply targets was smaller; from 20,000 in Nepal and 40,000 in Bangladesh to 200,000 in Pakistan and Niger and 228,000 in Madagascar (later increased to 278,000). Sanitation and hygiene results therefore dominated the programme overall, reflecting the focus of the Business Case - though in Myanmar, South Sudan and Yemen their targets for water supply were higher than those for sanitation. In most cases, ASWA constituted the bulk of the UNICEF country WASH programme.

2.4.2 Equity and Sustainability in Programme Design

The ASWA Business Case was clear that achieving equity of access was a priority concern and that the extent of inequity in access to WASH services should be one of the criteria for country selection. Beyond this, however, the document did not spell out how the programme should pursue equity. Neither did the logframe beyond adding 'including the lowest wealth quintile' to the indicator definitions for water and sanitation results at outcome level. The Business Case nevertheless indicated that the programme evaluation would place emphasis on 'equity and UNICEF's success in reaching out to the poorest in the most under-served areas.' Section 4.3 considers the extent to which COs met DFID expectations in this area.

The Business Case also highlighted the need for measures to enable the sustainable use of WASH facilities and the sustainability of hygienic behaviour. The logframe anticipated at least 5 million people gaining 'sustainable' access to WASH services in off-track countries. Beyond this, however, specific objectives and expectations relating to sustainability were not spelled out.

2.4.3 Budget Distribution

Table 4 shows how the budget has been distributed globally and expenditure as at June 30, 2016. 83% of the total was allocated to COs, while UNICEF headquarters and regional offices received roughly 10% and 4% respectively (the remainder was unallocated).

Of note in Table 4 is the expenditure of funds by the Madagascar CO which is 150% of budget. In fact, an additional USD $3,000,000 was allocated to Madagascar by UNICEF HQ in order to maximise the impact of ASWA investments and to reach overall beneficiary targets. However, the additional funds, allocated well before the date of the report summarised in Table 4, were not updated in the overall

[37] For details of targets see Table 4, Section 2.5 below.

financial data provided by UNICEF. The additional funds received and utilised were updated in the Madagascar Annual Report published in June 2016. Funds were used appropriately by the Madagascar CO. The financial data as of June 30, 2016 provided by UNICEF HQ had not been updated.

Table 4: ASWA Programme Financial Profile June 30 2016

Region	Country	Allocation USD	Total USD Expended/committed 30 June 2016	Cost-budget Ratio
Country Offices				
ROSA	Bangladesh	6,000,000	5,616,203	94%
	Pakistan	6,000,000	5,758,787	96%
	Nepal	5,000,000	4,484,683	90%
	Region total	17,000,000	15,859,673	93%
EAPRO	Myanmar	2,000,000	1,582,061	79%
	Cambodia	1,200,000	887,028	74%
	Region total	3,575,000	2,469,089	69%
ESARO	Madagascar	6,000,000	9,001,411	150%
	South Sudan	4,000,000	3,448,590	86%
	Region total	11,000,000	12,450,001	113%
WCARO	Niger	7,800,000	7,495,435	96%
	Region total	7,800,000	7,495,435	96%
MENA	Yemen	5,000,000	5,000,000	100%
	Region total	5,000,000	5,000,000	100%
Total Country Offices		**44,375,000**	**43,274,198**	**98%**
Regional and Head Offices				
ESARO RO		1,000,000	1,000,000	100%
WCARO		500,000	500,000	100%
EAPRO		375,000	375,000	100%
ROSA		0	0	
HQ		5,000,000	3,326,050	67%
Total Regional/Head Offices		**6,875,000**	**5,201,050**	**76%**
Total Allocated[i]		51,250,000		98%
Total Expended[ii]		48,475,248		
Unexpended[iii]		1,774,886		
Total Funds Expended & Programmed[iv]		50,250,134		
Total DFID Funding[v]		51,149,886		
DFID Funding Not Yet Designated[vi]		899,752		

Source: Annex P Programmatic VFM Study, pp.14-15. Notes to Table 4: [i]Total allocated to ASWA by UNICEF HQ; [ii] Total expended by ASWA to June 30 2016; [iii] Allocated funds not yet expended [iv] Total funds expended and unexpended; [v] Total DFID funds; [vi] DFID funds not yet allocated or expended

2.4.4 Theory of Change

As mentioned briefly above (Section 1.3), owing to limitations with the original TOC, the evaluation team in collaboration with DFID and UNICEF developed an elaborated ASWA TOC, which is set out in detail in Annex B. The elaborated TOC is a theoretical depiction of the logical relationships between the inputs, outputs, outcomes and impacts of the ASWA programme envisaged by the DFID ASWA Business Case and Logframe. It also highlights the key linkages (thirteen in total) and assumptions (sixteen in total) embedded in these relationships, as well as several rival explanations that could also contribute to the observed effects of the ASWA programme (i.e. factors outside of the programme such as other WASH programmes, changing contextual factors).

It is important to note that the original TOC was not used operationally to a significant extent by UNICEF in the design or management of the ASWA programme at global or country level. Rather, the logframe was the main operational framework. As such, the evaluation did not seek to assess the extent to which UNICEF used theory of change as a design and management approach. Rather, the evaluation sought

to use the elaborated TOC as an analytical framework (or hypothesis) by which to highlight and understand the revealed programme logic that emerged through the implementation of ASWA.
This meant that it was necessary to introduce and validate both the purpose and content of the elaborated TOC with UNICEF COs (e.g. producing country specific versions of it in collaboration with UNICEF COs for use in the Theory of Change Review method workshops) and other ASWA stakeholders during evidence collection for the evaluation. This introduced a somewhat artificial aspect to discussions of the elaborate ASWA TOC. This was mitigated by explaining to evaluation stakeholders that its use by the evaluation was primarily as an analytical framework and presented for them (what was often a welcome) opportunity to increase their own awareness of theory of change as a design and management approach.

2.5 Results to Date

The original target date for delivery of output level results was March 2016, but the programme was subsequently extended: firstly, on a no-cost basis and related to original targets, followed by a costed extension with additional outputs in five countries and a revised end date of March 2018. For the purposes of the evaluation, June 2016 was taken as the cut-off date for results data (see Section 1.5 above for rationale).

ASWA programme cumulative results for Outputs 1-4 are presented below. **Table 5** presents results as reported by UNICEF as of June 2016. These results were revised by IPME following output verification and the Sustainability Assessment Myanmar (SAM) and are presented in **Table 6**. The revised results are used by the evaluation in our analysis. The justification for revisions made are as follows.

Programme Results: Output 2.2: Adjusted total is based on data from country reports. However, this indicator has been effectively dropped by UNICEF COs and IPME due to a lack of clarity on what it means. Output 3.1: Rough estimate only, in the absence of accurate data for South Sudan.

Madagascar: Based on output verification, Output 1.1 is reduced but Output 1.2 and Output 3.2 are increased. The CO office has accepted these changes.

Myanmar: Output 1.1: Six monthly report states that 144 out of 215 triggered villages have become ODF, with a total population of 122,100. However, the SAM found that in a sample of 44 triggered villages, of which 28 had been reported as ODF, only 3 were actually ODF, and CLTS activities were already completed at the time of the survey. We have used the same ratio of 3/44 ODF to estimate that 15 out of 215 targeted villages were actually ODF by June 2016, with an estimated population of 12,719 (based on 15/144 x 122,100). We recognise, however, that this figure may later have risen, since in June 2016 the IPs were re-appointed for an additional 3-month input. Output 2.1: UNICEF/ DRD Myanmar assumed that the entire population of each target village gained access to water, but the SAM found that it was only 88%; we have reduced the figure accordingly. Output 3.2: The CO does not track the indicator for hand washing facilities, and the reasons for this remain unclear.

South Sudan: Output 1.2: Based on output verification, the figure has been reduced. Output 3.1: The cumulative total of 465,204 is unreliable as it consists largely of people reached via radio broadcasts. The 2015 appraisal flagged this error and the CO now only reports people reached through interpersonal communication every six months, but it has not amended the cumulative total. Output 3.2: Output verification found no documentary evidence to support the reported figure of 74,030.

Yemen: IPME cannot comment on any of the reported results as we have no access to source data and could not conduct a monitoring systems appraisal or output verification exercise due to security constraints. We were also unable to obtain any relevant data from third party monitoring agencies.

Niger: Following output verification, Outputs 1.1, 1.2; 3.1, 3.2 and 4.2 have been reduced to match IP reports.

Table 5: Cumulative results for Outputs 1-4 as reported by UNICEF in June 2016

	Madagascar	South Sudan	Niger	Cambodia	Myanmar	Pakistan	Bangladesh	Nepal	Yemen	Programme Results (Jun16)	Programme Target (Mar16)
Sanitation											
1.1	1,117,839	82,804	412,311	114,448	122,100	1,342,278	895,085	1,180,127	196,816	**5,463,808**	5,000,000
1.2	1,585,349	132,791	624,825	112,833	37,472	723,994	553,425	327,043	196,816	**4,294,548**	1,500,000
Water Supply											
2.1	304,170	242,250	200,250	0	98,957	215,570	40,500	24,294	152,652	**1,278,643**	900,000
2.2	10	18	176	0	20	0	52	68	56	**425**	576
Hygiene											
3.1	2,679,865	465,204	1,189,479	201,965	198,422	1,743,722	1,069,669	137,226	714,661	**8,400,213**	5,000,000
3.2	1,130,767	74,030	874,759	235,101	0	391,788	418,667	124,935	57,113	**3,307,160**	2,500,000
School WASH											
4.1	86	55	83	146	0	285	580	210	55	**1500**	120
4.2	341,495	34,287	56,741	58,000	0	37,852	172,632	127,942	32,000	**860,949**	120,000

Table 6: Cumulative results for Outputs 1-4 as revised by IPME following output verification and Sustainability Assessment, Myanmar

	Madagascar	South Sudan	Niger	Cambodia	Myanmar	Pakistan	Bangladesh	Nepal[38]	Yemen[39]	Programme Results (Jun16)	Programme Target (Mar16)
Sanitation											
1.1	939,078	82,804	342,713	114,448	12,719	1,342,278	895,085	1,180,127	196,816	**5,106,068**	5,000,000
1.2	1,546,764	118,363	555,247	112,833	37,472	723,994	553,425	327,043	196,816	**4,171,957**	1,500,000
Water Supply											
2.1	304,170	242,250	200,250	0	87,100	215,570	40,500	24,294	152,652	**1,278,643**	900,000
2.2	10	35	1	0	0	0	0	3	0	**49**	576
Hygiene											
3.1	2,679,865	?	1,071,570	201,965	198,422	1,743,722	1,069,669	137,226	714,661	**8,000,000+**	5,000,000
3.2	1,359,464	?	874,700	235,101	0	391,788	418,667	124,935	57,113	**3,307,160**	2,500,000
School WASH											
4.1	86	55	83	146	0	285	580	210	55	**1500**	120
4.2	341,495	34,287	50,000	58,000	0	37,852	172,632	127,942	32,000	**860,949**	120,000

[38] No output verification has been conducted so far in Nepal, due to the re-scheduling of IPME activities following the 2015 earthquake. It is planned for 2017.

[39] The IPME team has been unable to visit Yemen for security reasons, hence there has been no monitoring systems appraisal or output verification.

Comparisons of country level targets and revised results for Outputs 1-4 are presented in **Table 7** below. **Looking across the nine country programmes, most have met or exceeded their output targets, and by a considerable margin in some cases.** Madagascar and Pakistan, for example, both had ambitious targets from the start and, in the case of Madagascar, received additional funding which enabled them to raise their targets in 2015 and again in 2016. Both programmes over-delivered substantially for all the outputs. Similarly, in South Sudan and Cambodia results were almost double the target for water supply and hygiene promotion respectively. Yemen also shows substantial over-achievement on sanitation, but we cannot comment on the validity of any results from this country as the IPME team was unable to visit in person due to the ongoing conflict. **Where there are shortfalls, these are mostly quite small.** The exception is the ODF result for Myanmar, as the findings of the Sustainability Assessment Myanmar (SAM) led IPME to substantially reduce the figure reported by UNICEF.

Results under Output 5 (Enabling Environment) are all qualitative and in the absence of country-specific objectives or targets (except in Bangladesh) it is very difficult to confirm to what extent the output has been met. In many cases, it is not clear to what extent the activities reported were funded under ASWA. Results achievement for Output 5 is however discussed in Section 3.2.8[40].

[40] For detailed description of qualitative results under Output 5 Enabling Environment see Annex J, Section 1, Table 4.

Table 7: Comparison of country level targets and revised results[41] for Outputs 1-4 (Part 1)

Output	Output Indicator	Madagascar		South Sudan		Niger		Cambodia	
		Target (Mar16)	Results (Jun16)	Target (Mar16)	Results (Jun16)	Target (Mar16)	Results (Jun16)	Target (Mar16)	Results (Jun16)
Sanitation									
5 million people live in Open Defecation Free Communities, of which 1.5 million gain access to sustainable basic sanitation facilities	1.1 People living in Open Defecation Free communities	500,000[1]	939,078	42,000	82,804	400,000	342,713	120,000	114,448
	1.2 Number of people gaining access to sanitation as a direct result of the project	970,000[2]	1,546,764	42,000	118,363	400,000	555,247	120,000	112,833
Water Supply									
0.9 million people will gain access to sustainable improved water supplies	2.1 People gaining access to improved water supplies as a direct result of the project	278,000[3]	304,170	126,000	242,250	200,000	200,250	-	0
	2.2 Water resource and climate change assessments undertaken at watershed level prior to construction of water supplies	-	10	10	35	-	1	-	0
Hygiene									
10 million will be reached by hygiene education programmes of which 4.5 million people adopt habitual Hand Washing with Water and Soap (HWWS)	3.1 People reached with hygiene education programmes	970,000[4]	2,679,865	500,000	?[6]	1,000,000	1,071,570	100,000	201,965
	3.2 People with water and soap/ash available near their toilet (as per MICS/DHS proxy indicator for hygiene)	500,000[5]	1,359,464	100,000	?[7]	800,000	874,700	-[8]	235,101
School WASH									
400,000 school children will benefit from improved WASH facilities in schools	4.1 Schools with access to WASH (as defined by national standards)	50	86	50	55	70	83	270	146
	4.2 Number of children that have access to soap and water at school to practice HWWS	60,000	341,495	40,000	34,287	50,000	50,000	100,000	58,000

Notes

Boxes shaded green indicate over-achievement of target by at least 20%. Boxes shaded orange indicate under-achievement by at least 20%.

Madagascar

[1] Original target of 300,000 was increased when additional funding was received in 2015 and again in 2016

[2] Original target 500,000

[3] Original target 228,000

[4] Original target 500,000

South Sudan

[6,7] IPME Output verification found the numbers reported by UNICEF to be inaccurate, but revised figures could not be obtained

[8] While no target was adopted at programme start, results were delivered. The CO began reporting results in early 2016.

[41] Results as revised by IPME following output verification and Sustainability Assessment, Myanmar (see Section 2.5 above).

[5] Original target 300,000

Table 7: Comparison of country level targets and revised results for Outputs 1-4 (Part 2)

Output	Output Indicator	Myanmar		Pakistan		Bangladesh		Nepal	
		Target (Mar16)	Results (Jun16)	Target (Mar16)	Results (Jun16)	Target (Mar16)	Results (Jun16)	Target (Mar16)	Results (Jun16)
Sanitation									
5 million people live in Open Defecation Free Communities, of which 1.5 million gain access to sustainable basic sanitation facilities	1.1 People living in Open Defecation Free communities	104,963	12,719	870,000[9]	1,342,278	600,000	895,085	1,200,000	1,180,127
	1.2 Number of people gaining access to sanitation as a direct result of the project	24,560	37,472	652,500[10]	723,994	300,000	553,425	20,000	327,043
Water Supply									
0.9 million people will gain access to sustainable improved water supplies	2.1 People gaining access to improved water supplies as a direct result of the project	107,000	87,100	200,000	215,570	40,000	40,500	20,000	24,294
	2.2 Water resource and climate change assessments undertaken at watershed level prior to construction of water supplies	0	0	0	0	0	0	3	3
Hygiene									
10 million will be reached by hygiene education programmes of which 4.5 million people adopt habitual Hand Washing with Water and Soap (HWWS)	3.1 People reached with hygiene education programmes	165,200	198,422	1,050,000	1,743,722	1,000,000	1,069,669	189,000	137,226
	3.2 People with water and soap/ash available near their toilet (as per MICS/DHS proxy indicator for hygiene)	58,000	No data	146,000[11]	391,788	400,000	418,667	151,000	124,935
School WASH									
400,000 school children will benefit from improved WASH facilities in schools	4.1 Schools with access to WASH (as defined by national standards)	0	0	265	285	500	580	158	210
	4.2 Number of children that have access to soap and water at school to practice HWWS	0	0	36,000[12]	37,852	125,000	172,632	189,000	127,942

Notes

Pakistan

[9,10] Original targets 1.45m

[11] Original target 690,000

[12] Original target 21,600

Table 7: Comparison of country level targets and revised results for Outputs 1-4 (Part 3)

Output	Output Indicator	Yemen	
		Target (Mar16)	Results (Jun16)
Sanitation			
5 million people live in Open Defecation Free Communities, of which 1.5 million gain access to sustainable basic sanitation facilities	1.1 People living in Open Defecation Free communities	57,000	196,816
	1.2 Number of people gaining access to sanitation as a direct result of the project	57,000	196,816
Water Supply			
0.9 million people will gain access to sustainable improved water supplies	2.1 People gaining access to improved water supplies as a direct result of the project	130,000	152,652
	2.2 Water resource and climate change assessments undertaken at watershed level prior to construction of water supplies	0	0
Hygiene			
10 million will be reached by hygiene education programmes of which 4.5 million people adopt habitual Hand Washing with Water and Soap (HWWS)	3.1 People reached with hygiene education programmes	680,000	714,661
	3.2 People with water and soap/ash available near their toilet (as per MICS/DHS proxy indicator for hygiene)	57,000	57,113
School WASH			
400,000 school children will benefit from improved WASH facilities in schools	4.1 Schools with access to WASH (as defined by national standards)	50	55
	4.2 Number of children that have access to soap and water at school to practice HWWS	30,000	32,000

3. FINDINGS AND CONCLUSIONS

3.1 Introduction

Findings and conclusions are presented by ASWA output or theme (e.g. equity, VFM, programme monitoring) instead of by evaluation question. This has been done to enhance the utility of the evaluation report to the primary audience of WASH specialists in DFID and UNICEF. Within the sub-sections, however, the evaluation questions addressed are first set out in a table and then referenced in the text. A full summary of findings against evaluation questions is provided in Annex A.

At the end of each sub-section, a guide to evidence supporting the findings and conclusions is presented in a table to guide the reader. Each table presents the evidence sources for the evaluation questions (annexes, documentary evidence, specific case studies, or other source); the triangulation that has taken place using those evidence sources; and the strength of evidence. The strength of evidence is presented as per the categorisation presented in Table 8.

Table 8: Guide to Evidence

Strong	Strong (indicating corroborating evidence from multiple case study countries and ASWA programme globally)
Medium	Medium (indicating corroborating evidence either from multiple case study countries, OR ASWA programme globally, OR two or more of: documentary evidence; single case study; Other source).
Limited	Limited (indicating only one of: documentary evidence; single case study; other source).

Findings and conclusions related to the elaborated ASWA TOC are provided in Appendix 1 to this report and referenced in the text.

The rest of this section is structured as follows:

- Results achievement (Section 3.2)
- Quality of outputs (Section 3.3)
- Outcomes and sustainability (Section 3.4)
- Value for money (Section 3.5)
- Equity (Section 3.6)
- Programme monitoring (Section 3.7)

3.2 Results Achievement (EQ 2a, 2b, 3, 3a, 3b, 4, 4a, 4b, 7c, 11)

Table 9: Evaluation Questions

Evaluation questions addressed in this section	
2a	What were the key enabling and constraining factors to delivery of outputs and likely achievement of outcomes at the level of UNICEF (e.g. partnership arrangements and management, procurement strategy), implementing partners and communities / households?
2b	To what extent have outputs been delivered and outcomes likely to be achieved?
3	To what extent are donors and other development partners willing to coordinate and engage in joint learning to improve the WASH sector at national level?
3a	How effectively is UNICEF engaging with government to strengthen WASH sector coordination at national level?
3b	Has ASWA leveraged additional public financing to sustainably support strategic gaps in the WASH sector at all levels?
4	To what extent has ASWA improved enabling environment for WASH at national and / or sub-national levels and what has been its approach (e.g. in terms of sector systems and monitoring, strengthening government led scale up efforts, local government promotion / incentives and regulation, evidence based policy, and clarifying roles and responsibilities at decentralised levels)?
4a	How is government leadership and political will influenced to improve the WASH sector?
4b	To what extent is there local and national government capacity in key areas of sustainable WASH implementation?

7c	To what extent is the local private sector able to function profitably and willing to engage (e.g. in supplying services to hygiene promotion and / or operation and management of water supply)?
11	Have natural disasters or conflict prevented or impaired project implementation or damaged water or sanitation facilities post construction?

The focus of this section is primarily on how and why results were achieved at output level. In considering what might account for variations in achievement at output level across the nine countries programmes, it is important first to acknowledge some important differences between the country programmes, as outlined below**.**

3.2.1 Country contexts

3.2.1.1 Emergencies (EQ 11)

Many ASWA countries have been affected by national emergencies during the ASWA programme, not least Nepal which suffered a major earthquake in April 2015. This was followed by a protracted fuel blockade which brought programme operations to a virtual standstill for some months. Elsewhere, South Sudan and Niger were both affected by cholera outbreaks and conflict, while the outbreak of civil war in Yemen caused the country programme focus to shift its focus to emergency response. Though not a full-on emergency, the political situation in Bangladesh became volatile in late 2014 following elections earlier that year, with extended strikes and transport blockades. Meanwhile, Myanmar and Madagascar are both recovering from major political turmoil. The Myanmar country programme retains a large humanitarian WASH component, though ASWA operations have not been directly affected by emergencies.

3.2.1.2 Sector Status

The policy and institutional framework for WASH varies widely across the nine ASWA countries. It is difficult to pinpoint specific features which enabled or constrained the delivery of programme outputs, but it is evident that in the three South Asian countries the WASH sector is relatively mature compared to the other countries, in the sense that there has been considerable experience with community-based approaches to rural water supply and sanitation;[42]the Non-Governmental Organisation (NGO) sector is vibrant; and each country has to some extent established a national approach to sanitation promotion. This put the COs in a good position to operate programmes at scale. Cambodia also has a long history of sector support but has yet to see sanitation programmes operating on a large scale, while in Madagascar, Myanmar, Yemen, Niger and South Sudan sector institutions are weaker and less experienced in community-based WASH.

3.2.2 Country Programme Design

While ASWA represented what was essentially 'business as usual' for UNICEF, there were significant differences between the country programmes that had implications for the achievement of results.

3.2.2.1 Scale of Operations

The global programme design envisaged a total of five million beneficiaries for sanitation and 0.9 million for water supply. Most countries followed a similar pattern by prioritising sanitation over water supply, but the scale of operations varied considerably. Nepal had a sanitation target some 28 times greater than that in South Sudan and 21 times that in Yemen.

3.2.2.2 Implementing Partners

Most COs opted to work via NGO partners, at least for sanitation, hygiene promotion and school WASH. Where government played a direct role in implementation, this tended to be for water supply, particularly schemes with some degree of technical complexity.

[42] Bangladesh is the country where CLTS was first developed.

Madagascar stands out as the one country programme that delivered results, at scale, entirely through the government framework. The country has a dearth of capable national NGOs and suffers from weak government. UNICEF's response was to deploy its own consultants into government agencies at regional and commune (district) level. The Cambodia programme also works exclusively though government, but on a smaller scale and with less direct technical and operational support. The Niger programme initially tried to implement entirely through government but later brought in NGOs when it became clear that government alone would be unable to deliver in the time available.

3.2.2.3 New Versus Established Operational Models

UNICEF's approach to rural sanitation promotion known as Community Approaches to Total Sanitation (CATS) is based on CLTS, which is widely used in the sector globally. CATS is well-established within UNICEF, nevertheless each country office participating in ASWA has to customise their promotional methods and implementation strategy to suit local circumstances. The extent to which country programmes were ready to deliver at scale varied and this was reflected in the wide range of targets adopted.

3.2.3 Country Programme Management

3.2.3.1 Administrative Challenges

Though the programme officially started in August 2013, most COs did not receive funds until 2014. Consequently, field operations did not begin in earnest before late 2014 or early 2015. The implementation period available up to the original target date of March 2016 was, therefore, much shorter than originally envisaged.

3.2.3.2 Human resource constraints

Six out of nine COs have undergone a change in WASH Chief since the programme began. Periodic staff rotation is normal in UNICEF but nevertheless has implications for programme continuity, especially when there is a gap between outgoing and incoming managers. Probably the worst-affected country was Cambodia, where the incumbent WASH Chief left shortly after ASWA began and the post remained vacant for an extended period before the current Chief took up his post in mid-2015. In Myanmar, too, this post was vacant for a considerable period but UNICEF was at least able to appoint a very experienced caretaker manager.

The fact that country targets, whether large or small, were mostly met suggests that each country office adopted work plans and levels of ambition that were realistic under local circumstances. Based on the results achieved, there was little evidence of the 'optimism bias' that leads some grant-funded programmes to promise donors more than they can deliver.

Table 10: Guide to Evidence

Evidence Sources	Cross-Country Case Study Analysis (Annex H): EQs 2a, 2b, 7c, 11 Global Analysis (Annex I): EQs 2a, 2b, 11 IPME ongoing engagement with country offices: EQs 2a, 2b, 7c, 11
Triangulation	ASWA programme globally
Strength of Evidence	Strong: Corroborating evidence from multiple case study countries and ASWA programme globally

Based on this contextual analysis, it is possible to draw out some key trends and contrasts across the ASWA countries. These are discussed in detail in the following chapter.

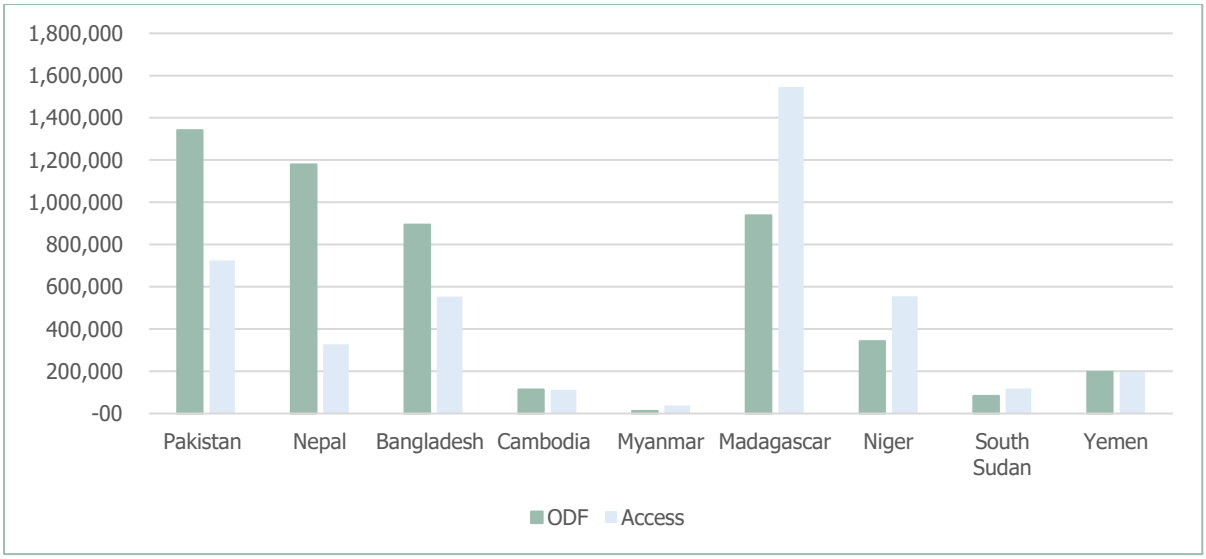
3.2.4 Sanitation

3.2.4.1 Enabling and constraining factors to the delivery of sanitation outputs (EQ 2a, 2b)

A note of explanation is needed on the relationship between the verified results for sanitation indicator 1.1 (people living in ODF communities) and 1.2 (people gaining access to sanitation) as presented in Table 4 (Section 2.4.1 above) and Figure 1 below. The revised logframe envisaged that the programme would result in an additional 5 million people living in ODF communities, of which 1.5 million would gain access to sanitation. In practice, not all people gaining access have come from ODF communities and the result for indicator 1.2 (4,171,957 beneficiaries) is only slightly lower than that for 1.1 (5,106,068 beneficiaries).

Furthermore, while four countries (Cambodia, Pakistan, Bangladesh and Nepal) have higher results for 1.1 than for 1.2, the reverse applies in four others (Madagascar, South Sudan, Niger and Myanmar); the (unverified) results reported for Yemen are equal. The reasons for this appear to be as follows. In some countries, such as Madagascar, baseline access to sanitation was close to zero in many of the targeted communities. After CLTS triggering, many toilets were built but only a minority of communities went on to become ODF. Consequently, the results for people gaining access were much higher than for those living in ODF communities. In the South Asian countries, however, the baseline level of access was already significant in many communities and so the effort required to get a community 'over the line' to ODF was less. Consequently, Pakistan, Nepal and Bangladesh all have much higher results for ODF than for people gaining access to sanitation.

Figure 1: Comparison of Sanitation Output Results by Country



Note: The IPME team has had no opportunity to validate the results reported from Yemen

The countries delivering the highest results for sanitation were Nepal, Pakistan, Bangladesh and Madagascar. **The South Asian country programmes all set ambitious sanitation targets, reflecting their long experience of CLTS, its adoption by government as a national approach and a relatively stable operating context.** It is nevertheless impressive that Nepal fell only slightly short of meeting its target despite there being two major crises (described above) during the programme.

Nepal stands out among the nine ASWA countries as having not just a national strategy for rural sanitation but a rural sanitation movement with strong government leadership at district level. This is especially so in the Terai region (the densely-populated plains along the Indian border) where much of ASWA activity takes place. Here, UNICEF is just one among several government and external agencies engaged in a concerted, co-ordinated effort to achieve district-wide ODF status. The initiative has considerable momentum and high visibility at local level.

Madagascar was something of an outlier in that it also adopted very ambitious sanitation targets (970,000) despite the weak sector framework, extreme poverty and acute water supply problems which would potentially make sanitation a low priority for rural communities. Here the key factors underlying programme success included an operational model that was tested and improved before ASWA began, and which featured institutional arrangements whereby UNICEF directly addressed human resource gaps in government agencies, giving them a high degree of control over activities on the ground.

In Cambodia, both UNICEF and government partners had been working with CATS / CLTS for some time before ASWA began, but programme ambitions were quite modest in this case. This was partly because of human resource constraints in government and an internal UNICEF requirement for the WASH programme to operate in the same locations as the country programme overall. This resulted in operations being spread quite thinly over a wide area, which was challenging for supervision and monitoring.

The Myanmar CO and NGO partners had limited prior experience with CLTS, and earlier projects had not operated at scale. **The evaluation mission found that the programme was still developing an effective approach and implementing partners needed further support and guidance. The Niger, South Sudan and Yemen teams were somewhere in the middle in terms of experience; operations in South Sudan and Yemen were also constrained by conflict.**

Several COs cited the limited technical capacity of implementing partners as a constraint in addition to the time taken to contract and orient them. In Niger, for example, UNICEF set out to implement the programme entirely via government agencies. When it became clear that the targets would not be met in the time available, UNICEF supplemented government capacity by contracting several national and international NGOs. This helped, but the NGOs, too, struggled to meet their targets until UNICEF intensified its technical support and guidance. In Pakistan, UNICEF had some 14 NGO partners of which some had substantial WASH experience and were professionally managed, while others were weaker both technically and in terms of management and administration. In Myanmar, the two NGOs implementing the CLTS component both had limited experience in this area. The duration of Project Cooperation Agreements (PCAs) with NGO partners was also a factor here. While twelve-month contracts were common, Pakistan extended them to 18 months. In Myanmar, however, they were just eight months. This gave IPs a very short time within which to deliver outputs and provide post-ODF follow-up in support of outcomes and sustainability (the long-term use and maintenance of toilets community-wide).

Given that country level sanitation targets were mostly exceeded, it appears that NGO capacity was ultimately not a constraint on results in most countries. Having said this, the extent to which COs included the development of NGO capacity in their operational plans and budgets was generally quite limited. UNICEF typically provided orientation and (where necessary) basic training at programme start on the implementation process and what was expected in terms of deliverables and reporting. In Pakistan, UNICEF also facilitated peer mentoring among the partner NGOs. COs did not, however, make it part of their strategy to provide comprehensive support to help create strong and effective local organisations that could serve as a sector resource in the longer term. Arguably, this is a gap that has implications for the sustainability of programme benefits. We note here that in Myanmar, where NGO capacity and experience with CATS was limited, the SAM found that the number of ODF communities was substantially lower than had been reported (see justification for IPME revisions to reported results in Section 2.5 above).

A further challenge identified by some respondents (both UNICEF and IPs) in some of the countries that lacked a standard national approach to sanitation promotion (for example, Cambodia and Myanmar) was that other sanitation projects operating in nearby locations were providing hardware subsidies. This was reported to undermine efforts to generate demand without subsidies under ASWA.

A summary of findings and conclusions on the programme logic from the elaborated ASWA TOC that are related to Output 1 can be found in Appendix 1, Table D.

Table 11: Guide to Evidence

Evidence Sources	Cross-Country Case Study Analysis (Annex H): EQs 2a, 2b, 11 Global Analysis (Annex I): EQs 2a, 2b, 11 Other: IPME ongoing engagement with country offices: EQs 2a, 2b, 11
Triangulation	Multiple case study countries (Myanmar, Pakistan, Madagascar) and ASWA programme globally
Strength of Evidence	Strong: Corroborating evidence from multiple case study countries and ASWA programme globally

3.2.5 Hygiene promotion including WASH in Schools

3.2.5.1 Enabling and constraining factors to the delivery of hygiene outputs (EQ 2a, 2b)

It is difficult to draw out comparisons across all nine ASWA programme countries on the hygiene promotion results and associated enabling or constraining factors underlying their delivery, for two reasons. **Firstly, no common understanding was established among programme stakeholders at the outset as to what it meant to be 'reached' with hygiene education programmes as per indicator 3.1.** Consequently, each country office followed their own course and in some cases no definition was ever clarified. Activities typically included: a mix of hygiene messages during CLTS triggering or follow-up meetings; one-to-one hygiene promotion via household visits; and other public and/or small group hygiene promotion meetings (for example, meetings for mothers of young children, meetings with men and boys). UNICEF thereby sought to ensure that the entire population of each targeted village was 'reached' in some way. In some cases, the hygiene promotion component was organised very systematically with activities described in detail in an operational manual. In Pakistan, for example, four hygiene promotion campaigns were undertaken in each village over a twelve-month period, with each one covering a different topic and having its own, customised Information Education Communications (IEC) materials (such as posters and flipcharts). **Regardless of problems with the indicators, the impression gained for the programme is that the hygiene promotion interventions were substantive and had a strong focus on hand washing in line with indicator 3.2 (people with water and soap / ash near their toilet).**

It is also noted here that the Cambodia programme set no target for hand washing facilities despite promoting hand washing under the programme. It only began reporting against Output 3 in early 2016. The Myanmar programme, meanwhile, adopted a target for indicator 3.1 but has never reported any results against it. The country office was unable to offer a clear explanation for this, though a lack of ASWA-specific baseline data was evidently a factor.

A second problem with interpreting the results is that for indicator 3.2 (people with water and soap/ash) it is impossible to discern the level of ambition of the programme since the target is not expressed as an improvement over baseline. **This aside, it is interesting to note that having a hand washing facility close to the toilet was one of the ODF criteria applied in some countries. This linkage could be viewed as an incentive for implementing partners and/or communities to intensify hand washing promotion. However, it also sets the bar for ODF quite high and consequently could depress ODF results.**

Three other enablers for the installation of hand washing facilities are evident:

- The ready availability in most locations of water for hand washing (which need not be potable), though there were seasonal shortages in some places.
- Promotion of very simple hand washing technology such as the 'tippy-tap' as used in Madagascar. This requires nothing more than an empty plastic bottle, some string and a few sticks, which can be periodically repaired or replaced by the household.
- The acceptance of ash (which is widely available) as an alternative to soap where this is expensive or not locally available.

Given the difficulties in defining and measuring people reached by hygiene promotion, and the fact that hygiene promotion is in any case an input rather than an output, indicator 3.1 is, arguably, of limited value. We note that DFID has already acknowledged this and does not plan to include it in the results framework for the successor programme to ASWA. The promotion of hand washing remains important as part of a package designed to reduce water and sanitation-related disease in the long term. However, a focus on outputs and outcomes rather than inputs would arguably be more useful as a management tool, though outcome measurement is challenging.

3.2.5.2 WASH in Schools

Given that UNICEF's mandate focuses on children, WASH in schools is an organisational priority. That ASWA targets for school WASH were quite modest was partly a reflection of DFID's commitment to deliver 60 million additional water and sanitation beneficiaries and the fact that school WASH results could not be counted against this target as part of the related MDG. Apart from this, some COs had already earmarked other funding for school WASH hence it was not a top priority for ASWA specifically. The bulk of programme results came from the three South Asian countries, while the Myanmar programme had no school WASH component at all.

One important enabler for school WASH interventions has been the application of UNICEF's 'Three Star' Approach which includes three levels of certification depending on the package of support offered to schools and the resulting level of WASH services and routine hygiene promotion activities. The Three Star package includes both hardware and 'soft' components: WASH facilities are brought up to national standard; systems put in place for their cleaning and maintenance; and teachers trained in providing hygiene education on an ongoing basis as part of the regular school timetable. This includes the addition of group handwashing at defined times. The more basic One Star certification stage has only the 'soft' elements and is hence much cheaper to apply at scale, though it may include the provision of water filters and simple hand washing facilities (as done in Madagascar). The promotion of mass hand washing at set times is increasingly being adopted as part of the schools' package in many UNICEF country programmes (within and beyond ASWA) and appears to be popular with students.

Under ASWA, the schools component has typically been implemented by the same partner staff who led sanitation promotion in the wider community, the intention being that school and community-based interventions should be mutually supportive. The extent of UNICEF's direct support to schools depends on the funding and human resources available. **In addition, most COs have engaged with government at policy level to pursue the introduction of policies and strategies to bring all schools up to an acceptable standard in terms of both WASH facilities and hygiene education.** This is a long-term initiative and the state of progress varies widely across the nine countries. UNICEF and its implementing partners report this to have been a productive strategy.

A summary of findings on the programme logic from the elaborated ASWA TOC that are related to Outputs 3 and 4 can be found in Appendix 1, Table E.

3.2.6 Sanitation and Hygiene

3.2.6.1 Enabling and constraining factors to the likely achievement of sanitation and hygiene outcomes (EQ 2a, 2b)

WASH Sector experience globally with rural sanitation and hygiene promotion shows that communities typically need further promotional inputs and long term advisory support and monitoring even after ODF has been achieved, to:

- consolidate latrine use by all family members;
- continue promoting hand washing with soap (changing behaviour is not easy and DFID regards even a 10% improvement over baseline as an acceptable result);
- ensure that damaged or full latrines are repaired, emptied or replaced; and
- ensure that new toilets are built as the population expands.

With this in mind, the following comments can be made on enabling and constraining factors for the delivery of programme outcomes relating to sanitation and hygiene.

3.2.6.2 Constraints

Once COs had received funds and procured their IPs, the implementation period remaining was much shorter than programme documents suggested and in most cases the PCAs were only long enough for the delivery of outputs. Most NGO partners were contracted for just one year and in Myanmar, they were contracted for only eight months (though they were brought back in for an additional three months later on). Pakistan was an exception; here contracts were for 18 months to allow for some degree of post-ODF follow-up. In Myanmar, time constraints were compounded by three factors. Firstly, the IPs had limited previous experience with CLTS. Secondly, there was no system in place for independent ODF verification. Thirdly, UNICEF struggled to provide adequate oversight to the sanitation field work following the departure of their Mandalay-based sanitation officer.

Since IPs are normally contracted for only for the programme implementation phase, post-ODF community support and monitoring needs to be provided by government agencies with a permanent presence in the local area. This could be a health or water department, or local government. **In the three case study countries, UNICEF was engaged in advocacy and technical support to help local agencies take on these responsibilities. But this was a work in progress and there was no guarantee that the necessary human and financial resources would be deployed, despite in-principle agreements from some government partners at local level.** This situation is typical of many low-income countries and resolving it is a long-term challenge for the sector. There are no easy answers but UNICEF and other development partners ensure that it features in sector debate at policy level as part of wider initiatives to foster sustainability. It is closely related to the subject of decentralisation, which again, is a work in progress in many countries.

There was no requirement under ASWA for toilets to meet Joint Monitoring Programme (JMP) criteria for 'improved' facilities, only that they should meet national standards. Not all countries had national standards and the quality of the facilities varied between countries and sometimes between regions within countries. In most of South and South-East Asia there is a strong community preference for flush toilets, irrespective of national standards, but these are not always built with durable superstructures. In Sub-Saharan Africa, meanwhile, dry pit latrines are more common and are often built with locally available materials that again are not durable. IPME country visits from 2014 onwards revealed that some UNICEF country programmes including Nepal, Bangladesh and Pakistan have (or have previously had) sanitation marketing initiatives. These promote and facilitate easy access to skilled labour and materials for the construction of affordable but durable and hygienic latrines in rural areas via the local private sector. However, in sparsely populated

rural areas such as Southern Madagascar, where there is extreme poverty and a very limited supply of sanitation goods and services, the scope for such market-based approaches is very limited. In such locations, toilets built with locally available (often natural) materials can function reasonably well, but may not withstand heavy winds or rains and may therefore need to be repaired or replaced seasonally. This is not a problem in principle, but sector experience from similar situations across Sub-Saharan Africa suggests that replacement might not happen without some ongoing motivational stimulus from government or external agencies. UNICEF country teams understand this and are trying to address it through advocacy and technical support. However, as indicated above, the challenge is huge as development agencies have little control over government actions post-implementation.

3.2.6.3 Enabling factors

Some COs applied operational approaches to sanitation and hygiene promotion that had been tested and improved prior to ASWA (for example in Pakistan and Madagascar). Here, the prospects for a lasting impact on toilet use and hand washing practices were likely to be better than those in places where the implementation approach was relatively new and the country office was still fine tuning it (as in Myanmar). See case studies (Annex E and G) for further discussion on how the promotional approaches used in Pakistan and Madagascar have evolved in recent years.

Where there is a strong, government-led initiative to promote rural sanitation and hygiene, outcomes and sustainability are less dependent on the direct contributions of UNICEF or other external agencies. Nepal stands out as an example of good practice in this respect; here UNICEF is one of many agencies supporting a co-ordinated multi-stakeholder campaign. In both Pakistan and Nepal, the adoption of a common strategy by government and development partners has led to operational synergies and activity has scaled up considerably. This has made it possible to target the achievement of entire ODF districts. Active government leadership and the scale of activity both contribute to the creation of a critical mass of promotional effort that could help to shift social norms around sanitation and hygiene. In Madagascar, ASWA is implemented entirely through government and UNICEF assistance is helping to develop technical and operational capacity to sustain ODF and promote hand washing. Although for now there is heavy dependence on UNICEF's own consultants deployed into government agencies. If funding ends and these personnel are no longer deployed, it is not clear what will happen (see Annex E Madagascar Case Study).

Table 12: Guide to Evidence

Evidence Sources	Cross-Country Case Study Analysis (Annex H): EQs 2a, 2b, 11 Global Analysis (Annex I): EQs 2a, 2b, 11 Other: IPME ongoing engagement with country offices: EQs 2a, 2b, 11
Triangulation	Multiple case study countries (Myanmar, Pakistan, Madagascar) and ASWA programme globally
Strength of Evidence	Strong: Corroborating evidence from multiple case study countries and ASWA programme globally

3.2.7 Water supply

3.2.7.1 Enabling and constraining factors to the delivery of water supply outputs (2a, 2b)

South Sudan and Yemen were the only two countries reported as off-track against the MDG target for water supply when the programme was planned. In most countries, ASWA targets and results for water supply indicator 2.1 (people gaining access to improved water supplies) were much smaller than those for sanitation, Niger and Yemen being the exceptions. There was no water supply component to ASWA in Cambodia.

It should be noted here that indicator 2.1 (water resource and climate change assessments undertaken at watershed level prior to construction of water supplies) was in effect dropped by UNICEF and DFID in mid-2015 due to a lack of clarity as to what exactly should to be done, and for which types of water

supply improvement. At the ASWA global meeting in April 2015, UNICEF headquarters staff proposed three possible levels of assessment and there was in-principle agreement by COs to adopt the most basic. Since then, however, most have not reported any results against this indicator. One explanation for this is that the proposal from headquarters was made rather late, when water supply improvements were already underway in most countries. It is also significant that the indicator in this case relates to inputs not outputs, and meeting the target would not have contributed to beneficiary numbers. For this reason, the indicator received less attention than other outputs at country and global level, although DFID continued to urge UNICEF to issue operational guidance.

Water supply improvements were predominantly made in communities that were also targeted for sanitation and hygiene. In Pakistan and Yemen, they were offered as a reward for achieving ODF status. In Madagascar, too, UNICEF's intention was to install a water supply scheme only after a community had become ODF. This condition was not applied in every case, however, due to acute water stress in some locations.

In Nepal, the IPME inception visit in 2014 found that UNICEF had prioritised support to government water supply schemes that had been started but not finished, and much of this work was in different locations to the sanitation and hygiene component of ASWA. Meanwhile in Bangladesh, where access to improved water supplies was already high[43] in rural areas, UNICEF used DFID support to improve access to the public water supply network in a slum in the capital. A member of the IPME team visited this sub-project in 2014.

UNICEF progress reports show that the bulk of water supply provision and improvement under ASWA has been in the form of communal point sources, typically dug wells and boreholes fitted with handpumps and operated under a community management model. **In a few cases, however, small piped schemes were also developed and in Myanmar and Madagascar there was only limited sector experience with the chosen service delivery model (which in Madagascar involved private operators) prior to ASWA.**

There have been a few instances of NGO partners delivering rural water supply improvements under ASWA (for example some of the NGO partners in Pakistan) but in most cases this work was led by government agencies. For the piped schemes in Madagascar and Myanmar, UNICEF provided technical support and quality assurance.

The installation of point sources was relatively straightforward both technically and in the sense that community demand for water is invariably high, hence the outputs were not predicated on behaviour change.

Bangladesh was something of a special case in that the water supply component of ASWA was a completely separate intervention to the rural CATS project, and focussed on one Dhaka slum (Sattola). Here UNICEF partnered with Dhaka Water and Sanitation Agency, which in turn contracted a number of NGOs and Community Based Organisations (CBOs) to work with households and facilitate the conversion of illegal connections into legal, metered connections. Each was shared by a defined group of households and supplied safe, reliable water. The project benefited some 40,500 people. The initiative to legalise such connections was first introduced under an earlier programme (not led by UNICEF) and ASWA provided an opportunity to extend it further. Since completion of the Sattola project, the utility has rolled out the initiative to many other slums in the city.

[43] In its 2015 update, JMP reported 87% rural access to improved water supply in 2014.

3.2.7.2 Enabling and constraining factors to the likely achievement of water supply outcomes (EQ 2a, 2b)

As with sanitation, there is a need for long term monitoring and support for water supplies under community management, though with slightly different objectives, which are to:

- Ensure the local availability of affordable spare parts;
- Provide technical support and occasional refresher training in operations and maintenance;
- Provide (or facilitate access to) skilled technicians for complex repairs that cannot be carried out by the managing user group; and
- Promote the consumption of water from safe sources.

The issues and challenges in ensuring that local government agencies take on these responsibilities apply here just as they do for sanitation. Sustainability depends less on securing behaviour change among users, but there is the additional challenge of trying to ensure a local supply of essential spare parts via the private sector, which can be difficult in remote and sparsely populated areas since there may be insufficient sales to make it commercially viable for local retailers to stock them.

The evaluation missions provided only limited exposure to point sources since the Myanmar programme focused on piped schemes and in Madagascar, while the programme supported a range of water supply technologies, the development of piped schemes was a relatively new initiative and a focus of learning for UNICEF and government partners. In Pakistan, meanwhile, water supply overall was a relatively small part of the programme and was therefore not prioritised in evaluation field visits.

The piped schemes developed in Myanmar and Madagascar were in effect pilots and their cost-effectiveness and long-term viability were still being tested. In Madagascar, the programme encouraged private contractors to operate and maintain the schemes but in the South, it had been very difficult to attract private sector interest. A more enabling environment is needed in both Myanmar and Madagascar to address barriers to increased private sector participation, particularly in the areas of regulation, procurement and governments' own capacity to engage in public-private partnerships.

UNICEF was closely monitoring the use and performance of the schemes in Madagascar (both technically and managerially) to learn to what extent they were delivering real benefits to all and what could be done to enhance their sustainability. In Myanmar, however, there was less evidence of UNICEF doing this. This raises a concern that risks to outcomes and sustainability might not be identified and addressed. **There was evidence in both countries that the revenue generated might not be sufficient to cover operation and maintenance costs (see Annex E Madagascar Country Case Study and Annex F Myanmar Country Case Study).**

In several countries, UNICEF is helping government tackle sustainability as a strategic issue at national level, and the absence of fully functional sector monitoring systems providing reliable data on the quality and reliability of services is a common constraint. A ubiquitous challenge in the WASH sector is that, where databases exist in less developed countries, they typically take the form of inventories. In other words, they record that a water point or scheme exists, and perhaps whether it was functional at a certain point in time, but the data is not subsequently updated. In some countries (Cambodia, Nepal and Bangladesh, for example) UNICEF is helping government to establish more useful management information systems. Furthermore, in Madagascar and Pakistan, UNICEF has spearheaded the introduction of periodic sustainability checks to inform sector planning and (hopefully) resource allocation. UNICEF headquarters is promoting and supporting the introduction of sustainability checks by all country programmes. Although this remains optional at country level given UNICEF's heavily decentralised institutional structure. At the time of writing, only Pakistan and Madagascar have conducted at least one sustainability check. Bangladesh plans to

disseminate a national sustainability survey carried out by another organisation but may fund a new sustainability check in future. Some of the countries have indicated a general intention to conduct a sustainability check at some point but at the time of the evaluation we were unaware of any other concrete plans.

A summary of findings and conclusions on the programme logic for the elaborated ASWA TOC related to Output 2 can be found in Appendix 1, Table F

Table 13: Guide to Evidence

Evidence Sources	Cross-Country Case Study Analysis (Annex H): EQs 2a, 2b, 7c, 11 Global Analysis (Annex I): EQs 2a, 2b, 7c, 11 Other: IPME ongoing engagement with country offices: EQs 2a, 2b, 7c, 11
Triangulation	Multiple case study countries (Myanmar, Pakistan, Madagascar) and ASWA programme globally
Strength of Evidence	Strong: Corroborating evidence from multiple case study countries and ASWA programme globally

3.2.8 Enabling environment (EQ 3, 3a, 3b, 4, 4a, 4b)

Summaries of country programme activities and results under Output 5 are presented in Annex J (Section 1 Table 4). These are based largely on six monthly reports to DFID but supplemented with the IPME team's own knowledge obtained from engagement with the COs since 2014. In most cases, the reports do not clarify the extent to which reported activities were funded under ASWA. It is apparent, however, that some of the activities related to long-term initiatives that began before ASWA and continued, drawing on more than one funding source. DFID has confirmed that, in this case, the evaluation should review overall country programme contribution to achievements relevant to the enabling environment output indicators, whether or not the activities were funded by DFID. The following paragraphs review achievements across the programme for each sub-indicator of Output 5.

Output 5.1 – Number of countries with evidence of improved sector monitoring as a result of programme support, including (but not limited to) reporting on equity of access and the use of monitoring data to inform sector planning

There is insufficient evidence available to assess the extent to which sector monitoring has improved in the ASWA countries since 2013. However, in those countries where the IPME team has had some exposure to UNICEF work in this area[44] (excluding Yemen and Niger) **it is evident that the support has been strategically significant, well-targeted and highly collaborative.** While there has been some progress, the sector has a long way to go in most countries. and the case studies found that reporting equity of access remains quite limited. UNICEF COs need more guidance on good practice as it is not obvious what more needs to be done in relation to CLTS, given that it aims to deliver improvements in sanitation across entire communities. Global experience shows that establishing viable sector monitoring systems is a long-term initiative that rarely fits into the time frame of a single donor-funded project[45]. Even after logframe revisions, this sub-output was defined in rather broad terms, perhaps unavoidably since it was designed to cover nine country contexts.

[44] At the time of the evaluation, IPME had provided TA on sector monitoring to UNICEF in Madagascar and Nepal, and provided advice on the design of sector-wide sustainability checks in South Sudan and Pakistan. Support to additional countries was anticipated for 2017.

[45] See, for example, Schouten and Smits Eds. (2015) From Infrastructure to Services, Trends in monitoring sustainable water, sanitation, and hygiene services. IRC, Netherlands. http://www.ircwash.org/resources/infrastructure-services-trends-monitoring-sustainable-water-sanitation-and-hygiene.

Output 5.2 – Number of country programmes that, in collaboration with government and other development partners, pilot a methodology for the measurement of indirect beneficiaries of external support

UNICEF's draft methodology for the measurement of indirect beneficiaries was discussed at the global ASWA meeting in Bangkok in April 2015. However, there was consensus among participants that it was unworkable. It was felt that the indicators to be assessed were mostly qualitative and could not be measured objectively. There has been no piloting under ASWA, though we understand that UNICEF is still pursuing the challenge at global level. We also note that the inclusion of indirect beneficiaries had been requested by UNICEF at the design stage; it was not a DFID requirement. UNICEF and other international development agencies have a longstanding desire to measure the effectiveness of their advocacy and TA and this is becoming increasingly relevant as countries transition to middle-income status and the focus of external support shifts towards 'upstream' work to strengthen to the policy and institutional framework for service delivery. This remains a difficult issue, however, and there is, so far, no consensus in the sector on whether it can be done in a credible way. We acknowledge the potential benefits of a viable methodology, but question the value of including specific numbers of indirect beneficiaries in future logframes before that methodology has been established.

Output 5.3 – Number of countries with evidence that government-led efforts to scale up sanitation and/or hygiene promotion have been strengthened via programme support.

Not all COs have pursued explicit strategies in this area but it is evident from the evaluation missions and earlier IPME country visits that **UNICEF is helping government to scale up WASH interventions (particularly in sanitation) in several countries (e.g. Madagascar, Cambodia, Niger, and Nepal).** Moreover, in most countries UNICEF is an active participant in sector fora at national level and proactively supports the strengthening of sector planning and co-ordination processes.

Output 5.4 – Number of countries with evidence that programme support has addressed critical capacity gaps in lead sector institutions

The provision of technical training and operational guidance for government agencies at sub-national level has been a staple of UNICEF WASH programmes for decades. At national level, most country programmes - including those participating in ASWA - support sector planning, co-ordination, monitoring and the development of sector policies and strategies, both directly and through participation in sector fora such as Technical Working Groups and sector workshops. UNICEF 's capacity building support to lead institutions is provided under the umbrella of a long-term partnership that extends beyond the boundaries of specific time-bound projects, hence the lines are somewhat blurred between activities funded via ASWA and from other sources. From the case studies, and the IPME team's ongoing engagement with COs since 2014, it is evident that **UNICEF support has been substantial and appreciated by government partners. Having said this, COs are not, in most cases, providing their support on the basis of a systematic needs assessment and capacity building plan that identifies and targets 'critical gaps'.**

Output 5.5 – Country programmes with active operational research on stunting, hygiene promotion (hand washing with water and soap (HWWS), sustainability of services, social norms and time saving related to WASH interventions.

UNICEF's June 2016 global progress report refers to **centrally commissioned studies**, all of which are concerned with water supply. The thematic focus of the research was relevant to ASWA objectives and the reports give the impression that the research was designed and conducted professionally. Beyond that, the evaluation team has insufficient information to comment on what has been achieved and to what extent findings have been disseminated at country level. In the VFM section of the same

report, UNICEF reports that £841,215 of the programme budget has been allocated to Burkina Faso, presumably for research, but there is no reference to allocations for any other research. **The research commissioned by UNICEF HQ, while relevant to the UNICEF programme globally and the WASH sector in general, has been largely disconnected from the rest of the programme. The justification for including centrally-managed research in ASWA is therefore unclear.**

The same UNICEF global progress report also listed **research undertaken by UNICEF COs**. The items included reflect a general lack of clarity in the programme as to whether COs should be reporting anything against this indicator; and what counted as operational research rather than operational analysis. **Some of what has been reported is clearly not operational research, particularly the adoption of integrated approaches to implementation, development of organisational partnerships and training for operational staff.** There are also some research activities that were initiated before ASWA and funded from other sources. **This said, there have been some significant country level research initiatives under the umbrella of ASWA** (for example, on WASH and nutrition in Bangladesh).

A summary of findings on the programme logic for the elaborated ASWA TOC that are relevant to Output 5 can be found in Appendix 1, Table G.

Table 14: Guide to Evidence

Evidence Sources	Cross-Country Case Study Analysis (Annex H): EQs 3, 3a,3b, 4, 4a,4b Global Analysis (Annex I): EQs 3, 3a,3b, 4, 4a,4b IPME ongoing engagement with country offices: EQs 3, 3a,3b, 4, 4a,4b
Triangulation	Multiple case study countries (Myanmar, Pakistan, Madagascar) and ASWA programme globally
Strength of Evidence	Strong: Corroborating evidence from multiple case study countries and ASWA programme globally

3.3 Quality of Outputs (EQ1,1a,1b, EQ 6a, 6d, 6e)

Table 15: Evaluation Questions

Evaluation Questions addressed by this section	
1	Is the design of each ASWA country project relevant to and coordinated with the on-going or planned strategy, impacts, outcomes and outputs being provided by national government and other key development partners (including NGOs); and with DFID's and UNICEF's country strategy / plan?
1a	Were the needs of target populations identified and used to inform the design of ASWA at country level?
1b	How is UNICEF HQ (PD WASH) introducing guidance and technical assistance on climate change adaptation and WASH to ASWA country offices?
6a	How is UNICEF HQ (PD WASH) introducing guidance and technical assistance on climate change adaptation and WASH to ASWA country offices?
6d	What is the level of quality of construction of water supply and sanitation services (in communities and schools)?
6e	What is the level of quality of supervision of operational water supply services (in communities and schools) by community level WASH committees and authorities supervising water service providers?

3.3.1 Programme Relevance

Each ASWA country programme was designed as a sub-set of the broader UNICEF country WASH programme. UNICEF country programmes are invariably designed in close collaboration with national government partners with the explicit aim of supporting national development strategies and priorities. At the design stage, UNICEF HQ required each country office to show in their proposal how the planned ASWA interventions were aligned with the Intermediate Results set out in the Country Programme Action Plan. **ASWA was therefore directly relevant to UNICEF country strategies and, by implication, government strategy and plans.** One possible exception to this is that the relevance of CLTS to the Myanmar context is questionable, given that open defecation in rural areas is already rare at just 6%, according to JMP data. Those data have, however, been widely criticised as unreliable

by sector stakeholders including UNICEF, and the ASWA baseline Knowledge Attitudes and Practices (KAP) survey found that open defecation in the four programme townships was 22%. The implication is that UNICEF should keep the relevance of their implementation strategy under review to ensure that it responds to the actual obstacles to long term latrine use and maintenance. There might, for example, be a case for laying greater emphasis on sanitation marketing to ensure that the poorest rural communities and households can install affordable improved latrines.

UNICEF also collaborates closely with other international development agencies and under Output 5, many of the country programmes have supported initiatives to strengthen sector policy, strategy, planning, co-ordination or monitoring.

Regarding relevance to DFID, Bangladesh was the only one of the nine countries that had an ongoing bilateral programme with a WASH component, though some others received WASH humanitarian support. Alignment of ASWA with DFID country strategies was not, therefore, a concern.

3.3.2 Output quality

The evaluation mission (and IPME country visits prior to that) provided only limited opportunities to assess the quality of physical outputs, whether hard or soft as this would have demanded a considerable amount of the time available in country. **The water supply and sanitation facilities seen during the evaluation missions and earlier IPME field visits were generally of a good or reasonable quality, but the number was too small to draw any conclusions on the programme overall.** However, monitoring systems appraisals carried out by IPME for seven of the nine country programmes in 2015 examined whether programme monitoring systems featured "mechanisms to monitor the quality of outputs, both 'hard' (e.g. water points) and 'soft' (e.g. hygiene promotion) and to address any shortcomings identified." The appraisals found that five out of seven country programmes monitored output quality comprehensively, while two out of seven did so partially. Four out of seven monitored 'hard' and 'soft' outputs separately, while three out of seven partially met the criteria. Furthermore, most UNICEF COs had regular meetings with implementing partners to review progress and address any shortcomings identified.

The most comprehensive quality assurance system was found in Pakistan, where the CO had appointed an independent agency for third party field monitoring (TPFM). The role of the TPFM agency was to monitor the quality of implementation processes and outputs based on detailed criteria, and to validate implementing partners' progress reports. It was deployed on a full-time basis for the duration of the programme, but the amount of time spent on direct observation in the field was customised to each partner, based on a capacity assessment at programme start. This activity was supplementary to the routine supervisory and monitoring visits, and progress review meetings, which UNICEF field staff conduct in all countries, security permitting. The frequency and scope of field visits varies according the size of the programme, the human resources available and the type of partner, since UNICEF normally has more access to, and control over, contracted NGOs than it does with government partners. In Madagascar, UNICEF can exert strong control over the quality of outputs since its own consultants play a direct role in programme implementation at regional and commune level.

Among the three case study countries, there appeared to be some weaknesses in field supervision and monitoring in Myanmar, where the IPME Sustainability Assessment (see Annex O) found that open defecation was continuing in many communities reported as ODF and that many of the new piped water supply schemes had significant problems with bacteriological water quality[46]. Contributing factors here

[46] Logframe outputs and indicators make no reference to water quality, neither is there a water quality dimension to JMP definition of improved water supply. Water quality findings do not, therefore, affect programme results at output level.

included the capacity of implementing partners and a vacant post in the UNICEF team which made it very difficult to provide adequate field-based support and guidance (see Annex F for further details).

3.3.3 Other evaluation questions relating to quality

3.3.3.1 EQ 1a. Were the needs of target populations identified and used to inform the design of ASWA at country level?

Documentary review of ASWA country programme proposals across all nine countries and key informant interviews in the three case study countries suggest that the **ASWA design at country level was in most cases not informed by needs identification of specific target populations.** In some countries pre-existing, secondary data sources on general WASH needs at national level were used to inform the design (e.g. UNICEF-WHO JMP data, analyses in government WASH strategies, national survey data). There is some evidence from the case study countries (Pakistan and Madagascar) that more detailed and contemporary data was used after design to inform implementation (e.g. baselines or special studies). Evaluation missions and IPME monitoring systems appraisals found that **UNICEF country programmes targeted the most under-served districts and communities at the planning stage. Thereafter, however, most did not track whether they were effectively serving those most in need within these areas.**

3.3.3.2 EQ 1b. How is UNICEF HQ (PD WASH) introducing guidance and technical assistance on climate change adaptation and WASH to ASWA country offices?

In 2014, UNICEF headquarters issued a WASH Climate Resilient Development Strategic Framework, though the evaluation has no information on the extent to which this has (or has not) influenced ASWA country programmes in any way. At the ASWA global meeting in April 2015, UNICEF HQ presented some outline guidance on how COs could conduct climate change and water resource assessments under Output 2. Further guidance was issued to COs, but requiring a new skillset, additional resources, and time. **Few countries have conducted such assessments (one very recent exception being work in progress in Bangladesh) – this output has effectively been shelved by the COs.**

3.3.3.3 EQ 6a. How appropriate are the locations for water supply services?

Insufficient evidence is available to answer this question for the programme as a whole. **Uncorroborated evidence from sustainability assessments at the sector level in Pakistan and for ASWA locations in Myanmar suggest that the majority of water supplies were appropriately located from a water safety perspective.** The Madagascar sustainability assessment did not consider the safety of water supply locations. In the locations visited for the evaluation the siting seemed appropriate both in terms of user access and protection from contamination. Some programmes included the provision of shared taps close to households (for example Madagascar, Bangladesh) or even house connections (Myanmar) thereby offering a high level of service in terms of accessibility.

Table 16: Guide to Evidence

Evidence Sources	Cross-Country Case Study Analysis (Annex H): EQs 1, 1a, 1b, 6a, 6d, 6e Global Analysis (Annex I): EQs EQ1, 1a, 1b, 6a, 6d, 6e IPME ongoing engagement with country offices: EQs EQ1, 1a, 1b, 6a, 6d, 6e
Triangulation	Multiple case study countries (Myanmar, Pakistan, Madagascar) and ASWA programme globally
Strength of Evidence	Strong: Corroborating evidence from multiple case study countries and ASWA programme globally

3.4 Outcomes and Sustainability (EQ 2, 6, 6b, 6c, 7, 7a, 7b, 7d, 10, 12)

Table 17: Evaluation Questions

Evaluation Questions Addressed in this Section

2	To what extent and why are people (in communities and schools): Using and maintaining sanitation facilities? Adopting hand washing with soap / ash and water after defecation? Using and maintaining water services?
6	To what extent are the benefits of the ASWA programme likely to continue after DFID funding ceases?
6.b	How willing are communities to participate in and mobilise for ODF?
6.c	How willing are communities to sustain ODF?
7	What are the major factors and drivers influencing the likely achievement or non-achievement of sustainability of the ASWA objectives for Sanitation Facilities? Water services? Hygiene?
7.a	To what extent is finance in place to sustainably support WASH sector systems development beyond external programme funding?
7.b	How adequate is government support to decentralised administrative levels?
7.d	To what extent are investments in WASH sector learning, coordination and other intangible activities valued by government and development partners?
10	Are WASH programmes and / or humanitarian programmes with WASH components being delivered by other development partners in the same locations as ASWA with related outputs and outcomes?
12	Have natural, political, economic forces or pandemic disease let to the displacement of target populations?

3.4.1 Introduction

As noted in Section 1.2, **virtually no relevant outcome data (endline versus baseline) was available by 30 June 2016 for the majority of ASWA programme countries. There were no outcome targets set at country level and there was a lack of clarity at programme start as to what COs should do in terms of outcome assessment.** This to some extent reflects the fact that a priority for DFID was to meet its target of 60 million beneficiaries at output level; this may have contributed to a lack of attention to outcomes in programme monitoring and annual reviews.

The IPME team contacted most COs in early 2015 to clarify what outcome data they would produce, and when. The IPME Team Leader also made a presentation at the ASWA global meeting in March 2016 encouraging COs to collect outcome data and offering technical support if required. Most countries indicated that they would produce outcome data, though some of the planned surveys or studies would not be ASWA-specific and as such the extent to which they would shed light on ASWA results was uncertain. The IPME team subsequently provided some TA to support the design of the Pakistan Sustainability Check and some initial guidance on a potential sustainability check in Bangladesh.

By the end of 2016, some relevant surveys had been completed, but their usefulness to the evaluation was constrained by several factors:

1. Sustainability checks undertaken in Pakistan and Madagascar were sector-wide and not closely aligned with ASWA logframe indicators.
2. The Pakistan endline survey was labelled a KAP survey but in fact gathered only output (access) data.
3. The Bangladesh CO commissioned a KAP endline survey in 2016 which mirrored the baseline in 2015. The report became available in late 2016 and contained some useful data, though the authors noted that observational data indicated an 'exceptionally high rate of handwashing using soap' which was 'strongly suggestive of observer and respondent bias /reactivity in influencing the results.'
4. In December 2016, Myanmar CO shared some preliminary data from their endline survey, though the full report was not yet ready[47]. It included comparisons of baseline and endline for a range of indicators, but the data was problematic in that:
   a) While the baseline had taken a random sample of all communities (urban and rural) across the four programme townships, the endline sampled only from rural communities that had participated in ASWA.

[47] At the time of writing, the report is still being finalised.

b) The endline data for household access to piped water was much lower than expected: only 12% in one township even though the project model was based on providing universal (or near-universal) access to house taps. UNICEF explained that some water supply improvements were still ongoing at the time of the survey, but even allowing for this the very low levels of access reported suggest that some of the surveyed villages were ones that had been targeted for sanitation promotion but not water supply improvements.
c) The baseline and endline surveys were conducted by different firms using different methodologies.

Given this situation, the evaluation has insufficient evidence to assess outcome performance It is only possible to analyse the prospects for sustainable outcomes based on available information on the country programmes and their sector context. The analysis below draws largely from case study country qualitative evidence (from the evaluation methods used during the evaluation missions to the case study countries), supplemented by quantitative and qualitative evidence for those countries from: WASH sustainability assessments / checks for those countries (two of which were sector-wide and not necessarily representative of ASWA); and a draft endline KAP survey report for one case study country. It also draws on ASWA programme level output target results.

3.4.2 Sanitation (Output 1)

3.4.2.1 Extent to which and why people are using and maintaining sanitation facilities (EQ2)

There is some country specific evidence that **some people are using sanitation facilities, with evidence from one country (Myanmar) that many people are doing so. There is evidence from one country that increasing latrine access is associated with reductions in OD, but strong country-specific evidence that (as is recognised in the WASH sector globally[48]) increasing latrine access by itself is not eliminating OD.** Elimination also requires behaviour change (as planned by ASWA) that is sustained over time to adopt sole use of latrines for defecation. There is insufficient evidence from analysis across all nine ASWA countries to corroborate this because baselines and special studies to assess ASWA outcome indicators have not been planned and / or completed by UNICEF for all nine countries. There is insufficient evidence from the three case studies or from analysis across all nine ASWA countries to draw findings and conclusions on the extent to which people are maintaining sanitation facilities. For the Case Studies this is because Sustainability Assessments did not survey people's maintenance of sanitation facilities[49] and for analysis across all nine ASWA countries this is because baselines and special studies to assess ASWA outcome indicators have not been planned and / or completed by UNICEF for all nine countries.

3.4.2.2 Extent to which benefits are likely to continue after DFID funding ceases (EQ 6)

There is strong country-specific evidence that **ODF slippage is a real risk once implementing partner contracts end and post-ODF follow-up responsibilities fall on government agencies which require additional funding for staff and community actors in these roles**. UNICEF is working to address these issues, but the challenge is not yet resolved. There is corroborating evidence from analysis across all nine ASWA countries (but not output specific) for the above that potentially makes this finding externally valid (see Overall EQ6 findings and conclusions below Section 3.4.7.1).

Case study reports for Myanmar (Annex F, pp. 32-33) and Madagascar (Annex E, p. 23) highlight that a lack of government commitment to formalise post-ODF follow-up responsibilities within funded

[48] See, for example, http://www.communityledtotalsanitation.org/page/clts-approach.

[49] The Myanmar, Pakistan and Madagascar sustainability assessments / checks surveyed the enabling conditions for maintenance of sanitation facilities (e.g. institutional arrangements, availability of spares, user groups) but did not assess actual condition of sanitation facilities beyond how clean they were.

workplans is a factor (either because of funding constraints or a lack of agreement across government to prioritise funding for them).

3.4.2.3 Extent to which communities are willing to participate in, mobilise for and sustain ODF (EQ 6.b, 6.c)

There is strong country-specific evidence that **communities are willing to participate in and mobilise for ODF when CLTS methods are adapted based on evidence from national implementation that reflects contextual factors, including social norms and village size.** There is some country-specific evidence that the co-existence of subsidy approaches can be a constraint to community willingness to participate in and mobilise for ODF. There is strong country-specific evidence that **communities' willingness to sustain ODF has been improving, but can slip rapidly if there is a gap in post-ODF follow-up between the departure of implementing partners and take up of responsibilities for implementation of post-ODF activities by responsible government and community actors.** There is corroborating evidence from analysis across all nine ASWA countries (but not output-specific) for the above that potentially makes this finding externally valid (see Overall EQ6 findings and conclusions below Section 3.4.7.1).

3.4.2.4 Major factors and drivers influencing the likely achievement or non-achievement of sustainability of the ASWA objectives for sanitation facilities (EQ7)

There is strong country-specific evidence that the choice and implementation of strategies for the removal of barriers to the installation and sustained use of improved latrines need to be based on evidence about national contexts, including social norms and baseline levels of latrine access and OD. It should not be assumed that CLTS is appropriate in all national or sub-national contexts. For example, in countries where levels of latrine access are already high, sanitation marketing (with market / supply chain enabling environment support as needed) may be a valid approach in combination with, or instead of, CLTS. There is some country-specific evidence that sanitation monitoring and reporting systems in the WASH sector in ASWA countries (that should enable sustainability assessment) need strengthening, including ODF verification systems (see also Section 3.2.8 Output 5.1 above).

Table 18: Guide to Evidence

EQ 2	
Evidence Sources	Cross-Country Case Study Analysis (Annex H) EQ 2
Triangulation	Multiple case study countries: Pakistan, Madagascar, Myanmar
Strength of Evidence	Medium: Corroborating evidence from multiple case study countries
EQ 6	
Evidence Sources	Cross-Country Case Study Analysis (Annex H): EQ 6 Global Analysis (Annex I): EQ 6 Specific Case Study: Myanmar (Annex F, pp. 32-33, Section 4), Madagascar (Annex E, p. 23, Section 3.1.4)
Triangulation	Multiple case study countries: Pakistan, Madagascar, Myanmar ASWA programme globally
Strength of Evidence	Strong: Corroborating evidence from multiple case study countries and ASWA programme globally
EQ 6.b, 6.c	
Evidence Sources	Cross-Country Case Study Analysis (Annex H): EQ 6.b, 6.c Global Analysis (Annex I): EQ 6
Triangulation	Multiple case study countries: Pakistan, Madagascar, Myanmar ASWA programme globally
Strength of Evidence	Strong: Corroborating evidence from multiple case study countries and ASWA programme globally
EQ 7	
Evidence Sources	Cross-Country Case Study Analysis (Annex H): EQ 7
Triangulation	Multiple case study countries: Pakistan, Madagascar, Myanmar
Strength of Evidence	Medium: Corroborating evidence from multiple case study countries

A summary of findings on the programme logic from the elaborated ASWA TOC that are relevant to Output 1 can be found in Appendix 1, Table H.

3.4.3 Hygiene (Output 3)

3.4.3.1 Extent to which and why people are adopting HWWS after defecation (EQ2)

There is some country-specific evidence that **few people are adopting HWWS after defecation, with evidence from one country (Myanmar) that some people are doing so. There is some country-specific evidence of increases over baseline (of between 9-20% from baseline) in household investment in facilities that would enable them to practice HWWS** (which is in the middle range compared to evidence in other WASH evaluations reviewed: 13-46%)[50]. Although this is a proxy and not a direct indicator of handwashing practice, the increase is a positive result. There is insufficient evidence from analysis across all nine ASWA countries to corroborate the above because baselines and special studies to assess ASWA outcome indicators have not been planned and / or completed by UNICEF for all nine countries.

3.4.3.2 Extent to which benefits are likely to continue after DFID funding ceases (EQ 6)

There is strong country-specific evidence that **there are outstanding needs for HWWS awareness raising and follow-up and a lack of clarity on how this will be funded going forward**. There is complementing evidence from analysis across all nine ASWA countries (but not output specific) for the above that makes this finding potentially generalisable – see Overall Section 3.4.7.1 (EQ6) findings and conclusions below.

3.4.3.3 Major factors and drivers influencing the likely achievement or non-achievement of sustainability of the ASWA objectives for hygiene (EQ7)

There is insufficient evidence from the three case studies to draw findings and conclusions on major factors and drivers influencing the likely achievement or non-achievement of sustainability of the ASWA objectives for hygiene. This is because evidence at country level could not be corroborated across methods (so was not strong) as Rapid Outcome Assessment was the only method that in practice generated evidence for hygiene for this evaluation question across case study countries. However, as hygiene promotion was integrated with CLTS in all three case study countries, the findings and conclusions under EQ7 for sanitation are likely to apply to hygiene also (see Section 3.4.2.4 above).

Uncorroborated evidence from the Rapid Outcome Assessment method is presented for illustration only. It suggests that achievement of hygiene behaviour change was in part influenced by IPs effectively linking with diverse local government and civil society organisations who supported implementation (Myanmar and Madagascar). Analysis from one country case study (Pakistan) suggests that HWWS slippage after the departure of IPs is a risk, partly because the focus of sustainability is on ODF and partly because hand washing practices are notoriously difficult to assess.

A summary of findings on the programme logic for the elaborated ASWA TOC that are relevant to Output 3 can be found in Appendix 1, Table I.

Table 19: Guide to Evidence

EQ 2	
Evidence Sources	Cross-Country Case Study Analysis (Annex H): EQ 2
Triangulation	Multiple case study countries: Pakistan, Madagascar
Strength of Evidence	Medium: Corroborating evidence from multiple case study countries

[50] Other WASH evaluations reviewed showed evidence for increases in hand washing facilities as follows: 13% (AAN Associates (2014) Evaluation of the UNICEF Sanitation Programme at Scale in Pakistan (SPSP) – Phase 1, p. 51, Fig.24); 25% (e-Pact (2017) Evaluation of the WASH Results Programme Mid-Term Evaluation Report, p. 82); 46% (Contzen, N., Meili, I. H., et al. (2015). 'Changing Handwashing Behaviour in Southern Ethiopia: A longitudinal study on infrastructural and commitment interventions.' Social Science and Medicine 124: 103-114).

EQ 6	
Evidence Sources	Cross-Country Case Study Analysis (Annex H): EQ 6 Global Analysis (Annex I): EQ 6
Triangulation	Multiple case study countries: Pakistan, Madagascar, Myanmar ASWA programme globally
Strength of Evidence	Strong: Corroborating evidence from multiple case study countries and ASWA programme globally

3.4.4 School WASH (Output 4)

There is insufficient evidence from the three case studies or from analysis across all nine ASWA countries to draw findings and conclusions on: the extent to which, and why, people in schools are using and maintaining sanitation facilities (EQ2); and the extent to which benefits are likely to continue after DFID funding ceases (EQ6). This is because baselines and special studies to assess ASWA outcome indicators for school WASH have not been planned and / or completed by UNICEF for all nine countries. In addition, for the two case study countries that included the School WASH (Pakistan and Madagascar), the sector sustainability assessments in those countries did not cover school WASH.

Indicators for outcomes and prospects for sustainability of School WASH should be included in the design of future baselines and special studies to assess outcomes, and in sustainability assessments.

A summary of findings on the programme logic for the elaborated ASWA TOC that are relevant to Output 4 can be found in Appendix 1, Table J.

3.4.5 Water Supply (Output 2)

3.4.5.1 Extent to which and why people are using and maintaining water services (EQ2)

There is insufficient evidence on levels of use of water services. Water supply was only a major component of ASWA in two case study countries (Myanmar and Madagascar) and evidence on the level of use specific to ASWA is only available for one (Myanmar). There is some country-specific evidence that **choice is a factor in continued use of unsafe sources some of the time even when an improved source is available.** The choice to use unsafe sources can be for reasons of taste (Myanmar) or to only use safe sources at critical times (Madagascar). Further evidence is needed on continued use of unsafe water when improved sources are available. There is insufficient evidence from analysis across all nine ASWA countries to corroborate the above because baselines and special studies to assess ASWA outcome indicators have not been planned and / or completed by UNICEF for all nine countries. There is insufficient evidence from the three case studies or from analysis across all nine ASWA countries to draw findings and conclusions on the extent to which people are maintaining water services. For the Case Studies this is because Sustainability Assessments did not survey people's maintenance of water services and for analysis across all nine ASWA countries this is because baselines and special studies to assess ASWA outcome indicators have not been planned and / or completed by UNICEF for all nine countries.

3.4.5.2 Extent to which benefits are likely to continue after DFID funding ceases (EQ 6)

There is some country-specific evidence that **water systems may not be financially viable because the costs of operation and maintenance are not fully covered by tariff levels and revenues, which may themselves be constrained by willingness and / or ability to pay.** There is corroborating evidence from analysis across all nine ASWA countries (but not output specific) for the above (see Overall Section 3.4.7.1 EQ6 findings and conclusions below).

3.4.5.3 Major factors and drivers influencing the likely achievement or non-achievement of sustainability of the ASWA objectives for water services (EQ7)

There is insufficient evidence from the three case studies to draw generalisable findings and conclusions on the major factors and drivers influencing the likely achievement or non-achievement of sustainability of the ASWA objectives for water services. This is because findings and conclusions at cross-country study level could not be corroborated across countries (so was not generalisable) as Madagascar was the only case study country that in practice generated findings and conclusions for water services for this evaluation question across the case study countries.

One country case study (see Annex H, EQ7, Madagascar Findings and Conclusions) presents strong sector evidence that tariff payments that should fund operation and maintenance are low (33% of communities) and present a significant obstacle to sustainable water services. The situation is better for systems under private operation in Madagascar. However, given the limited ability and / or willingness of rural communities to pay tariffs and the additional costs of suppling more remote ones, establishing revenue policies that are financially viable and socially acceptable is a challenge for sustainability that is not yet resolved for community or private management in Madagascar. This evidence could not be triangulated for EQ7 with the other case studies.

A summary of findings on the programme logic for the elaborated ASWA TOC that are relevant to Output 2 can be found in Appendix 1, Table K.

Table 20: Guide to Evidence

EQ 2	
Evidence Sources	Cross-Country Case Study Analysis (Annex H): EQ 2
Triangulation	Multiple case study countries: Madagascar, Myanmar
Strength of Evidence	Medium: Corroborating evidence from Multiple case study Countries
EQ 6	
Evidence Sources	Cross-Country Case Study Analysis (Annex H): EQ 6 Global Analysis (Annex I): EQ 6
Triangulation	Multiple case study countries: Madagascar, Myanmar ASWA programme globally
Strength of Evidence	Strong: Corroborating evidence from multiple case study countries and ASWA programme globally

3.4.6 Enabling Environment (Output 5)

3.4.6.1 Extent to which finance is in place to sustainably support WASH sector systems development beyond external programme funding (EQ7a)

There is some country-specific evidence that **funding to sustainably support WASH sector systems (both institutional arrangements and infrastructure) beyond external programme funding does not appear to be in place.** WASH ministry ambition and plans to increase WASH finance are emerging but will require strong advocacy across government and with development partners to fill funding gaps. Case study reports for Madagascar (Annex E, Section 3.1, p. 19) and Myanmar (Annex F, Section 3.1, p. 17) highlight that UNICEF advocacy and technical advice is supporting government efforts to develop WASH strategies and frameworks that aim to increase sector finance.

3.4.6.2 Extent to which government support to decentralised administrative levels is adequate (EQ 7.b)

There is strong country-specific evidence that **government support to decentralised administrative systems is partially adequate, with needs for clarity in roles and responsibilities and their financing.** Case study reports for Madagascar (Annex E, Section 3.1, p. 20), Myanmar (Annex F, Section 3.1, p. 19) and Pakistan (Annex G, Section 3.1, p. 18) highlight that a lack of clarity in roles and responsibilities and their financing at decentralised levels can reflect silos and / or overlap between separate ministries that have stakes in WASH at national level.

3.4.6.3 Extent to which investments in WASH sector learning, co-ordination, and other intangible activities are valued by government and development partners (EQ 7.d)

There is strong country-specific evidence that **UNICEF investments in learning and coordination are valued by government and other development partners and UNICEF is seen as a leading contributor to such efforts.** Case study reports for Madagascar (Annex E, Section 3.1, p. 21) Myanmar (Annex F, Section 3.1, pp. 19-20) and Pakistan (Annex G, Section 3.1, pp. 18-19) highlight that UNICEF technical advice, facilitation and advocacy for and within sector groups has supported progress in policy environments that government and other development recognise as often being challenging and / or limiting for WASH sector co-ordination and learning.

Table 21: Guide to Evidence

EQ 7.a	
Evidence Sources	Cross-Country Case Study Analysis (Annex H): EQ 7.a Specific Case Study: Madagascar (Annex E, p. 19, Section 3.1), Myanmar (Annex F, p. 17, Section 3.1)
Triangulation	Multiple case study countries: Madagascar, Myanmar
Strength of Evidence	Medium: Corroborating evidence from multiple case study countries
EQ 7.b	
Evidence Sources	Cross-Country Case Study Analysis (Annex H): EQ 7.b Specific Case Study: Madagascar (Annex E, p. 20, Section 3.1), Myanmar (Annex F, p. 19, Section 3.1), Pakistan (Annex G, p. 18, Section 3.1)
Triangulation	Multiple case study countries: Madagascar, Myanmar, Pakistan
Strength of Evidence	Medium: Corroborating evidence from multiple case study countries
EQ 7.d	
Evidence Sources	Cross-Country Case Study Analysis (Annex H): EQ 7.d Specific Case Study: Madagascar (Annex E, p. 21, Section 3.1), Myanmar (Annex F, pp. 19-20, Section 3.1), Pakistan (Annex G, pp. 18-19, Section 3.1)
Triangulation	Multiple case study countries: Madagascar, Myanmar, Pakistan
Strength of Evidence	Medium: Corroborating evidence from multiple case study countries

3.4.7 Overall (not output specific)

3.4.7.1 Extent to which benefits are likely to continue after DFID funding ceases (EQ 6)

There is evidence from analysis across all nine ASWA countries[51] that almost all **COs' programmes, and UNICEF HQ, articulate and understand the need for support to sustain outputs after a community is mobilised or infrastructure is completed[52]. COs do plan for sustainability by ensuring that community management committees are established, and support from government structures is often promised for post-output support. However, many COs express concerns about the sustainability of WASH outputs.** Maintaining the progress made under ASWA creates a burden upon the next programme iteration to monitor and support the sustainability of WASH benefits. At this stage, it is not possible to state that sufficient resources have been dedicated within the implementation phase to support the beneficiaries of ASWA going forward. From a VFM perspective, funds explicitly dedicated to post-output support, and carefully expended have potential to increase sustained outputs.

UNICEF HQ has been encouraging COs to develop plans for sustainability checks and this was flagged heavily at the ASWA global meeting in April 2015. **By the end of 2015, only Madagascar and Pakistan had concrete country-specific sustainability plans.** Only five of nine ASWA countries requested IPME support for sustainability planning through the first quarter of 2016 (South Sudan, Nepal, Madagascar, Pakistan, and Myanmar).

[51] For further information on the PVFM evidence for EQ6 see Annex P Programmatic VFM Study.

[52] For further information see Programmatic VFM Study Annex P Section 2.9.4, pp. 27-30.

3.4.7.2 Extent to which other WASH programmes are present and constitute a plausible direct rival explanation for ASWA programme level outputs and outcomes (EQ10)

There is some country specific evidence that other development partners with related WASH outputs and outcomes are present in some ASWA locations, though on a smaller scale in most cases. Further evidence would be needed to assess the extent of the contribution / overlap in these countries. There is insufficient evidence from analysis across all nine ASWA countries to corroborate or complement these findings. The country case study report for Myanmar (Annex F, p. 55) highlights that in some ASWA village locations between one and three development partners were investing in inputs that contributed to the same outputs and outcomes as ASWA, but these were at a much smaller level and were not separate programmes. The country case study report for Pakistan (Annex G, p. 48) highlights that in some ASWA districts between one and two development partners had separate programmes with related WASH outputs and outcomes, but these were in a maximum of two of the eleven districts ASWA was located in.

Based on the above, no strong evidence has been found that would constitute a plausible direct rival explanation for ASWA programme level outcomes (i.e. where other development partners' programmes account for the results). WASH programmes and / or humanitarian programmes with WASH inputs at the same scale as ASWA are not being delivered by other development partners in many of the same locations as ASWA with related outputs and outcomes. There is some evidence that other development partners with related WASH outputs and outcomes from inputs at a smaller scale than ASWA are present in some ASWA locations. Further evidence would be needed to corroborate or reject whether this could constitute a plausible commingled rival explanation for any ASWA programme level outcomes (i.e. where ASWA and other development partners both contributed to the results).

3.4.7.3 Have natural, political, economic forces or pandemic disease led to the displacement of target populations (EQ 12)

Analysis across all nine ASWA countries does not reveal any evidence that insecurity, conflict, political instability, elections, natural disasters or epidemics have displaced people in ASWA locations (even when wider populations nationally may have been). There is insufficient evidence from the three case studies to corroborate this finding. UNICEF ASWA Global Annual Reviews (Annex J, GARQual) and Country Six-Monthly Reports (Annex J, 6MRQual) highlight that security, conflict, political instability, elections, natural disasters or epidemics were present in many ASWA countries (see also Section 3.2.1.1 above). However, the reported effects were to divert or impede the actions of UNICEF staff and its implementing partners and / or government counterparts rather than displace target populations from ASWA locations.

No strong evidence has been found for natural, economic forces, or pandemic disease having led to the displacement of target populations that would constitute a plausible super rival explanation for ASWA programme level outcomes (i.e. where a force majeure larger than ASWA accounts for the results).

Table 22: Guide to Evidence

EQ 6	
Evidence Sources	Global Analysis (Annex I): EQ 6
Triangulation	ASWA programme globally
Strength of Evidence	Medium: Corroborating evidence from ASWA programme globally
EQ 10	
Evidence Sources	Cross-Country Case Study Analysis (Annex H): EQ 10 Specific Case Study: Myanmar (Annex F, p. 55), Pakistan (Annex G, p. 48)
Triangulation	Multiple case study countries: Myanmar, Pakistan
Strength of Evidence	Medium: Corroborating evidence from multiple case study countries

EQ 12	
Evidence Sources	Global Analysis (Annex I): EQ 12 Other: UNICEF ASWA Global Annual Reviews (Annex J, GARQual) and Country Six-Monthly Reports (Annex J, 6MRQual)
Triangulation	ASWA programme globally
Strength of Evidence	Medium: Corroborating evidence from ASWA programme globally

3.5 Value for Money (EQ5,5a,5b)

Table 23: Evaluation Questions

Evaluation Questions Addressed in this Section	
5	Is the programme delivering Value for Money in terms of effectiveness, and equity and why?
5.a	What are the evidence streams to support the conclusions reached about Value for Money?
5.b	Is the programme delivering Value for Money in terms economy and efficiency and why?

The evaluation design envisaged a comprehensive Programme level VFM analysis and a more limited country level analysis. The evaluation has exceeded our initial plan and conducted country level VFM analyses for all nine countries[53], as well as the planned Programme Level Analysis. A significant limitation to the VFM analysis is the scarcity of outcome level data at the programme or country level. Outcome data is the evidentiary basis for effectiveness measures in VFM analysis. Effectiveness is a measure of the uptake and sustainability of tangible outputs, and the broad socio-economic benefits that are derived by beneficiaries. Outcome data requires resource-intensive surveys and other tools sometimes after outputs are completed, to ascertain beneficiary uptake and benefits. Less significant limitations to the VFM analysis and the extent of mitigation are noted in Annex P (Section 1.4.4).

The analysis is presented by theme. The findings and conclusions derive from evidence from programmatic VFM analysis across all nine ASWA countries. As such all findings and conclusions are generalisable to the ASWA programme (in terms of statistical generalisation) but have limited external validity (as there is no control group). There are no related elements of programme logic in the theory of change for this sub-section.

3.5.2 Effectiveness and Equity

3.5.2.1 Extent to which and why the programme is delivering value for money in terms of effectiveness and equity (EQ 5)

There is evidence from analysis across all nine ASWA countries for the following findings and conclusions. VFM in terms of effectiveness cannot be fully analysed as valid outcome level data was not available for the majority of ASWA countries or at programme level. In the absence of such data, **qualitative assessment of the potential of improved outcomes shows that programme level achievement has been above targets for key output indicators, which is a positive indicator of programme traction, and sets a strong foundation for sustained outcomes and effectiveness.** To strengthen the effectiveness of ASWA country programmes, DFID and UNICEF should allocate a proportion of funds specifically for follow-up to support the transition from outputs to outcomes within the funded implementation phase (for analysis of outcomes and sustainability see Section 3.4 above).

VFM in terms of equity cannot be fully assessed as gender disaggregation is the only measure for which data are available. Gender disaggregation data is estimated and largely follows general population trends. Thus, actual variations in accessing and use of benefits are assumed and these estimates serve little purpose. **Equity measures beyond gender estimates should be undertaken, especially where beneficiaries may be denied access to benefits due to caste, ethnicity, or cultural**

[53] For further information on country level VFM analysis see Annex P Section 3, pp. 31-54.

barriers. In a programme where considerable programme flexibility to shift funds across outputs and countries exists, it is imperative to measure equity of access further; at least by caste or ethnic grouping or by wealth quintile, to ensure that programme flexibility does not undermine output equity. IPME monitoring and verification component also highlights that some IPs in some countries do in fact have disaggregated household data on socio-economic status, ethnicity etc. but have not always reported if UNICEF hasn't asked for such data.

In planning and reporting, both **UNICEF and DFID should be aware of the potential negative impact on effectiveness and equity that may be caused by: a) the drive to maximise output results numbers; and b) the ongoing unmet burden of funding and support to sustain WASH Outputs after ASWA ends** (see also Section 3.6 Equity below).

3.5.3 Economy and Efficiency

3.5.3.1 Extent to which and why the programme is delivering value for money in terms of economy and efficiency (EQ 5.b)

There is evidence from analysis across all nine ASWA countries for the following findings and conclusions. **The programme is delivering VFM in terms of economy and efficiency, with some variation across countries**. Average ASWA programme unit costs are below those budgeted in the 2014 ASWA Business case for three of four key outputs[54]. Indirect to direct costs[55] are below 10% in four of nine ASWA countries (Bangladesh, Madagascar, Niger, and Pakistan) and below 15% in seven of nine countries (add Myanmar, Yemen, and South Sudan). Nepal and Cambodia are higher (see Table 24 below and full details in Annex P Section 4). Under 11% indirect costs demonstrates an apparent economy of scale in that UNICEF CO.

Table 24: Indirect to Direct Costs by Country

Country	Indirect to Direct Cost %
Madagascar	4.1
Niger	4.5
Pakistan	9.1
Bangladesh	9.5
South Sudan	11.3
Myanmar	13.9
Yemen	14.0
Nepal	20.0
Cambodia	30.6

For the programme as a whole they are under 11% (see Table 25 below). Comparing the percentage of direct programme level expenditures to indirect costs (programme support and monitoring and evaluation) is one measure for assessing economy and efficiency used in the Programmatic VFM Assessment. **Under 11% indirect costs demonstrates an apparent economy of scale provided by UNICEF.**

Table 25: Total Programme Direct to Indirect Costs

Outputs	Total by Output	% of Total
Sanitation	14,562,838	31.6%
Water supply	16,954,442	36.8%
Hygiene	2,813,771	6.1%
School WASH	4,036,106	8.8%
Enabling Environment	2,786,374	6.0%
Programme support and M&E	4,953,937	**10.7%**

[54] Comparing DFID ASWA Business Case budgeted unit costs to ASWA actual unit costs is one measure for assessing and benchmarking economy and efficiency used in the Programmatic VFM Assessment (see Annex P, p. 12).

[55] Direct costs are incurred in delivering programme outputs and indirect costs are incurred in managing programme delivery.

Total CO spend	46,107,468	

Procurement and budget management for each country programme is embedded within UNICEF's global systems and provides on-demand detailed financial analysis for programme managers, and clear metrics and procedures for economic procurement of commodities. In ASWA, most CO procurement is done by the IPs which may increase costs, but also greatly increases efficiency and timeliness of output delivery. From a VFM perspective, increased costs can be justified if the procurement strategy is timely and shifts risk away from UNICEF.

Key economy measures involve the procurement and risk mitigation of implementing partners. There is evidence that Madagascar, Bangladesh, Nepal, and Pakistan employ well established protocols within UNICEF to achieve procurement of economic and results-oriented partners. We were not able to assess partner selection data for Yemen or South Sudan. A primary focus of the Niger programme is capacitating government structures, creating less need for independent procurement. A limited pool of suppliers in Cambodia constrained partner selection. In Myanmar, the primary channel for outputs was through government and community structures and financial management was managed through Direct Cash Transfer (DCT) and Harmonized Approach to Cash Transfer (HACT) protocols.

Of note are the contracting models used by Madagascar to contract drillers which limits and shifts risk away from UNICEF and to the contractor. The Madagascar CO uses Long Term Agreements (LTAs) with a pool of contractors to enable agile response to needs without new contracting. Fixed price contracts for drilling a successful borehole in a community are established. Payment of the contract is by the result - a functional borehole of satisfactory quality. Failed boreholes are the driller's responsibility and cost. Using LTAs to prequalify a pool of suppliers, then hold the suppliers to a fixed price contract payable on results is a potential lesson to be shared across ASWA country programmes.

Also worth mentioning is the Results Allocation Framework (RAF) methodology used by the Nepal CO, which is an effective method to direct and budget funds through partners with agility for results. As part of its regular operating procedure the WASH section consults with partners, donors, and stakeholder to prepare an expected framework of agreed targets and available funds. A separate framework is created for each partner, enabling the WASH section to engage with active results management for results and efficient funds allocation within budgets. In environments where partners and other obstacles can retard results, the RAF is an approach to maximize results from inputs and minimize the risk of underutilized funds.

The agility to drive global results by fund reallocation to country programmes with capacity to scale quickly, is a high value VFM approach, maximising both economy and efficiency. At the same time, the reallocation of funds away from lesser achieving programmes may be rewarding already successful programmes. By reallocating funds for greater results, there is potential that funds will be moved away from difficult to reach populations, thus compromising equity which would undermine VFM. We encourage further beneficiary disaggregation by wealth quintile; caste, or another meaningful identifier. Despite this real tension, we view the reallocation of funds to achieve results as powering positive economy and efficiency.

While we note above that the reallocation of funds to maximize results can be seen to have contributed to increased economy and efficiency, strong and consistent monitoring support, and strategies to strengthen sustainability of WASH inputs are also vital. If the drive for beneficiary numbers incentivises programmes to concentrate on overachieving their output targets at the expense of focusing on longer-term outcomes, the effectiveness and VFM of initial investments could be undermined.

There is insufficient evidence to draw findings and conclusions on the use and results gained from funds expended by the Regional Offices. **To facilitate improved costing, DFID and UNICEF are encouraged to compile unit costs for key interventions across countries, ascertaining differences in the components for unit-costs so that a range of costs for standardised components by intervention is gradually developed.**

3.5.4 Evidence Streams

3.5.4.1 Evidence streams that support conclusions reached about value for money (EQ 5.a)

There is evidence from analysis across all nine ASWA countries for the following findings and conclusions:

Overall: The conclusions reached about ASWA Value for Money are supported by review, analysis, disaggregation, and triangulation of quantitative (financials and indicators) and qualitative data (UNICEF country reports, HQ procedures, and DFID approaches) from multiple sources. Primarily from the country level but complemented by HQ and Regional levels and IPME monitoring and verification data.

Economy and Efficiency VFM: At the country level, output results data appear to be credible for key outputs, though there was initially some misunderstanding around the meaning of output indicators 1.1 and 1.2 and how to count the results.

The degree to which available financial data are informative for VFM analysis is in question. Specifically, Country Office budgets appear to be regularly re-adjusted to meet actual expenditures if there is a significant variance from the planned budget. For example, in the 2016 global review some country budgets (Nepal, Pakistan, Myanmar, Cambodia and South Sudan) were changed downward, others were revised upward (Yemen and Madagascar) and two remained unchanged (Niger and Bangladesh). The inconsistency in reporting budget changes, which appears to bring actual expenditures closer to the "budget", may have been done at the CO level to keep budgets and expenditures "close". This may be a natural outgrowth of UNICEF's stringent monitoring of fund utilization leading to the reallocation of funds if they will not be used for one output, to another. This is understood, but best-practice financial reporting is to ensure that a record of budget changes is clear so that there is a trail of evidence to understand how and why changes were made. The rationale is that the comparison of the original budget and the actual expenditures "tells a story" about programme operations and context that caused expenditure changes. When the budget is adjusted to resemble the actual expenditure, the reasons and context for programme changes remain obscured. **While there is value in agile budget readjustment and reallocation of resources to other needs, at the same time budget readjustment to closely match expenditures does not permit VFM analysis to examine the efficiency of initial programme planning and budgeting.** The implication is that when budgets are substantially under or overestimated, the efficient use of funds is undermined, as is VFM[56].

Also at the country level, some indicators are not associated with disaggregated budget and expenditure data, the result of combining different types of activities and outputs (sanitation and hygiene education, for example). This further weakens VFM analysis of specific outputs.

[56] For further information see Annex P Section 2.7, p. 16

At the programme level, the strengths and weaknesses of Country Office performance and financial data are transferred upstream. **There does not appear to be significant quality assurance of CO data presented to the programme level**, though it is possible that the reports seen do not reflect the interactions between HQ and the COs to improve reported data if requested by HQ. It will be useful at the programme level to establish clear protocols for COs to follow when reporting financial data (budgets and expenditures) and performance data (targets and results). Such protocols should also set out how Indicators are defined.

Budget and expenditures for ROs are clear, but it is not clear how such funds were spent to support ASWA COs. UNICEF standardized frameworks for financial, results, and sustainability data capture and reporting should be designed, agreed upon, and communicated to recipient countries before the next phase of WASH programming commences with UNICEF. This is needed to ensure standardized application by COs when implementing programmes such as ASWA.

Effectiveness and Equity VFM: Programme-level assessment of the effectiveness of ASWA interventions is not possible because valid outcome level data was not available for the majority of ASWA countries. As further ASWA programming is developed and funded, **allocations to gather and assess outcome data from current ASWA programmes would strengthen future VFM analysis**.

Evidence to support equity measures in ASWA is more limited than desired. **Data capture to demonstrate the equity dimension of ASWA is not well-designed relying, largely, upon assumed gender disaggregation in the general population.** Additional equity measures (wealth quintile, caste, <5 beneficiary) would be useful to demonstrate equity in future programming.

Table 26: Guide to Evidence

EQ 5	
Evidence Sources	Global Analysis (Annex I): EQ 5
Triangulation	ASWA programme globally
Strength of Evidence	Medium: Corroborating evidence from ASWA programme globally
EQ 5.b	
Evidence Sources	Global Analysis (Annex I): EQ 5.b
Triangulation	ASWA programme globally
Strength of Evidence	Medium: Corroborating evidence from ASWA programme globally
EQ 5.a	
Evidence Sources	Global Analysis (Annex I): EQ 5.a
Triangulation	ASWA programme globally
Strength of Evidence	Medium: Corroborating evidence from ASWA programme globally

3.5.5 VFM Capability

The majority of UNICEF COs in ASWA have come to appreciate the evidence that VFM analysis generates, as an aid in making informed programme decisions. The challenge / opportunity is to identify and customize VFM support for each CO. The needs of each are different, but the majority are poised and eager to take VFM forward. **One constraint to more use of VFM by COs is the need to fund VFM support.**

3.6 Equity (EQ 8,8a,8b)

Table 27: Evaluation Questions

Evaluation Questions Addressed in this Section	
8	How appropriate was the selection of communities and schools in terms of targeting benefits at communities in the lowest wealth quintile and at women and girls?
8.a	To what extent has UNICEF's approach to promoting equity been operationalised and successful or not?

8.b	To what extent are processes and tools in place to identify, target and monitor equity between population groups with different social characteristics based on their needs?

Equity considerations are explicitly included in the DFID ASWA logframe (2015 – See Annex C) outcome indicators 1 and 2 in terms of the beneficiaries of water services and sanitation facilities including those in the lowest wealth quintile. No quantitative target is given to make explicit what would count as having achieved inclusion. However, the intention that benefits should include the poorest and certainly not exclude them is clear in the ASWA design. The evaluation was also asked to look at the gender dimension of equity (i.e. inclusion of women and girls), although this was not an indicator in the ASWA logframe. In the second half of the ASWA implementation period, DFID began to request 'disaggregated beneficiary data' as part of the DFID results returns but were not clear as to what this meant in practice for sanitation given that ODF means that the entire community has ended open defecation and started using toilets.

3.6.2 Targeting of Benefits

3.6.2.1 Whether the selection of communities and schools was appropriate in terms of targeting benefits at communities in the lowest wealth quintile and at women and girls (EQ 8)

The ASWA logframe from DFID did not provide a quantitative target to make explicit what inclusion of beneficiaries in the lowest wealth quintile for water services and sanitation facilities would look like. Neither did it include any type of indicator for gender. Greater clarity from the outset would have helped to drive equity planning and outcomes.

There is some country-specific evidence that **in terms of ASWA design, only a partially appropriate approach to targeting benefits at communities in the lowest wealth quintile was used by UNICEF (despite equity being formally mainstreamed in UNICEF's programming approach globally).** Partial because the approach only relates to water or sanitation benefits (when the ASWA logframe suggested it should relate to both water and sanitation benefits) and because it is insensitive as to whether it over- or under-includes the whole of the lowest wealth quintile.

There is strong country specific evidence that **UNICEF's approach to targeting benefits at communities by wealth presents two design risks.** Firstly, the risk of inclusion of wealthier households / communities and exclusion of poorer households / communities. Secondly, the risk of implementation errors not being systematically managed or lesson learnt due to no formal monitoring of equity outcomes. A design to benefit the whole community does not necessarily mean that outcomes are equitable unless targeting, baselines and monitoring of equity are used.

A summary of findings on the programme logic for the elaborated ASWA TOC that are relevant to Equity Targeting can be found in Appendix 1, Table L.

3.6.3 Approach to Promoting Equity

3.6.3.1 Extent to which UNICEF's approach to promoting equity been operationalised and successful or not (EQ 8.a)

There is strong country-specific evidence that **UNICEF's approach to achieve equity through community wide / universal benefits for WASH within geographically selected areas, without specific targeting by wealth quintiles or at women and girls, has been operationalised** (Myanmar, Pakistan and Madagascar). Some countries do, however, target water or sanitation subsidies at the poorest 5% or poorest 20 households within communities (Pakistan and Madagascar). There is complementing evidence from analysis across all nine ASWA countries that there is a general lack of information from UNICEF on the extent to which its approach to promoting equity

has been operationalised, and successful or not. This is because reporting on equity has not been mainstreamed within ASW.

A summary of findings on the programme logic for the elaborated ASWA TOC that are related to equity promotion can be found in Appendix 1, Table M.

3.6.4 Processes and Tools

3.6.4.1 Extent to which processes and tools are in place to identify, target and monitor equity between population groups with different social characteristics based on their needs? (EQ 8.b)

There is strong country-specific evidence that **UNICEF has guidance / plans and good practice / insights on gender and WASH (e.g. on identification, monitoring, and evaluation) which were not applied in ASWA and could have driven better equity planning and outcomes[57].** There is corroborating evidence from analysis across all nine ASWA countries that in many countries processes and tools are not in place under ASWA to identify, target and monitor equity between population groups with different social characteristics based on their needs (see also Section 3.7.2 Point 7 below). In a few countries, UNICEF does monitor aspects of equity, but even here this does not appear to be for all variables or ASWA locations. There is also complementing evidence from analysis across all nine ASWA countries that UNICEF HQ / RO has supported countries to better address and report on gender issues. However, **more needs to be done in this area, including support to COs on developing capacity in UNICEF good practice in equity identification, targeting and monitoring relevant to WASH.** Equity guidance to sanitation activities that aim to deliver improvements across entire communities using CLTS methods may require further research in the WASH sector.

Table 28: Guide to Evidence

EQ 8	
Evidence Sources	Cross-Country Case Study Analysis (Annex H): EQ 8 Specific Document: DFID ASWA Logframe (Annex C)
Triangulation	Multiple case study countries: Madagascar, Myanmar, Pakistan
Strength of Evidence	Medium: Corroborating evidence from multiple case study countries and Documentary Evidence
EQ 8.a	
Evidence Sources	Cross-Country Case Study Analysis (Annex H): EQ 8. a
Triangulation	Multiple case study countries: Madagascar, Myanmar, Pakistan
Strength of Evidence	Medium: Corroborating evidence from multiple case study countries
EQ 8.b	
Evidence Sources	Cross-Country Case Study Analysis (Annex H): EQ 8.b Global Analysis (Annex I): EQ 8. b
Triangulation	Multiple case study countries: Madagascar, Myanmar, Pakistan and ASWA programme globally
Strength of Evidence	Strong: Corroborating evidence from multiple case study countries and ASWA programme globally

3.7 Programme Monitoring

3.7.1 Introduction

This section addresses programme monitoring, which has been so central to the IPME assignment. IPME engagement in this area has generated many insights and lessons related to the evaluation overall. The evaluation included Evaluation Questions 7.e 'How effectively is UNICEF embedding community-led monitoring and the role of civil society in monitoring and accountability systems?'. However, there was insufficient evidence from analysis across the case study countries or across all

[57] For examples of relevant UNICEF CO gender and WASH documents see Annex E, p. 32, Annex F, p. 30, and Annex G, p. 27.

nine ASWA countries to draw findings and conclusions for this EQ. The rest of this section discusses programme monitoring within ASWA more broadly.

3.7.2 Output and Outcome Monitoring

One of the tasks identified in the IPME TOR was 'quality assuring programme progress reports, ensuring that robust baseline data is collected by the programme and assessing whether results reported are attributable to the programme.' The TOR were not specific as to how this should be done, but following discussions with DFID during the inception period, it was agreed in late 2014 that IPME would appoint national consultants, or 'Country Monitors' to carry out appraisals of programme monitoring systems used under ASWA. Thereafter, Country Monitors would conduct one output verification exercise per country[58] to validate the latest six-monthly report from UNICEF to DFID.

The relatively late decision to conduct appraisals, in addition to the time it took to recruit, orient and train the Country Monitors, and then to confirm timetables with UNICEF for carrying out their work, meant that the appraisals took place much later in the programme lifecycle than would be considered ideal. The bulk of the work was completed by July 2015 with Nepal completed somewhat later, having been postponed following the massive earthquake in April 2015. No appraisal was possible in Yemen due to the ongoing conflict.
A standard process and criteria for traffic light scoring were adopted for the appraisals. The process included 18 areas of investigation arranged in five broad categories: structure, functions and capabilities; data collection and reporting; data management processes; alignment with national reporting; and general. The appraisal focused primarily on the monitoring of field-based activities related to Outputs 1-4 in the ASWA logframe.

In general, programme monitoring systems were found to be fit for purpose. The appraisals nevertheless identified multiple issues and challenges including the following:

Output level

1. Most country programmes supplied IPs with proformas for monthly or quarterly reporting but provided only limited guidance on how the source data should be generated and managed.

2. Some IPs and UNICEF COs faced human resource constraints, limiting their ability to deploy full-time monitoring staff.

3. Several country programmes commissioned independent baseline studies but these were of limited value for output monitoring when based on limited samples. Community profiles produced in each village as part of the CATS process were more useful to UNICEF and IPs.

4. The weakest area of output monitoring was counting the number of people 'reached' by hygiene promotion (and avoiding double counting when the same people were reached through multiple activities).

5. Monitoring by government agencies of IPs' or their own UNICEF-funded activities was generally weak.

6. COs recognised the need to promote, enable and monitor the sustainability of programme results, but few had incorporated sustainability factors into routine monitoring. This said, an increasing number were supporting the introduction of sector-wide sustainability checks. On

[58] In the event, six output verification exercises were completed by the time of the evaluation missions.

the establishment of post-ODF monitoring and support, this remained a work in progress in most cases.

7. UNICEF country programmes targeted the most under-served districts and communities, but thereafter most did not track whether they were effectively serving those most in need within these areas. Most routine reporting did not disaggregate beneficiary data in terms of poverty, disability or gender.
8. COs made use of monitoring information in regular progress review meetings with implementing partners. However, reporting formats tended to focus mostly on quantitative results with less attention to the quality of processes and outputs.

9. Alignment with sector monitoring systems was a hypothetical issue in most ASWA countries as none had a system that was operational and effective nationwide (a common situation in less developed countries).
10. Few COs had established mechanisms for linking programme costs and results as part of programme monitoring, though VFM training and technical support from IPME was helping to address this.

Outcome level

11. Most UNICEF COs did not expect IPs to report outcome data and this is likely to be accounted for by:

    a) The programme timeframe being short and in nearly all countries, field work starting late. The focus of programme activity and monitoring was therefore on delivering the outputs which made the outcomes possible. Furthermore, no guidance had been issued on DFID expectations in this area.

    b) Project Co-operation Agreements (PCAs) tending to be short (typically 12 months) and ending with the delivery of outputs.

    c) Changes in the number of people using toilets and practising HWWS could not be measured through short visual checks as part of routine reporting. Special studies were required.

We noted earlier that there was a lack of clarity and consensus between DFID and UNICEF at programme start regarding outcome assessment. **Irrespective of a steer from DFID, it was incumbent on UNICEF to put arrangements in place for outcome assessment, given that the logframe included three outcome indicators**, albeit based on a simple assumption of 100% conversion of outputs. In the event, virtually no relevant outcome data (presented as progress against baseline in ASWA project locations) was available by 30 June 2016.

Two key factors contributed to the lack of outcome data. Firstly, many UNICEF COs routinely commission periodic KAP studies which investigate the use of WASH facilities and hygienic practices. However, these are not usually project-specific. More commonly they are done for the country programme overall and/or for the sector in general, often as part of the five-yearly programme planning cycle. At programme inception, the question of how and when COs would measure progress against the ASWA outcome indicators specifically was not given sufficient consideration by UNICEF headquarters and/or COs. Also, very few conducted the sort of baseline assessments that would later enable the measurement of quantified progress against outcome indicators. By the time the IPME team was appointed, the programme was already approaching its first annual review and it was too late to influence or support the design of baseline surveys, though the need to assess outcomes was subsequently flagged by DFID and the IPME team.

Secondly, both DFID and UNICEF were committed to meeting beneficiary targets at output level. The programme got off to a slow start, and by the time of the first annual review was significantly behind schedule. After DFID raised concerns about the rate of progress, UNICEF adopted measures to accelerate implementation and the targets were eventually met. There was no intention by UNICEF to marginalise outcome assessment, but several COs were still focussed on the delivery of outputs when the original target date of March 2016 was reached, and inevitably, the timing of outcome assessments was pushed back.

One other limitation of programme monitoring and reporting was that there was, on more than one occasion, a mismatch between the data and other information presented in country office six-monthly reports and the content of global reports from UNICEF headquarters. Some of the differences were due to mathematical errors at global level, but the reasons for other variances are unclear. For example, the very large number of water resource assessments appearing in global summary reports.

3.7.3 IPME Practice

While the evaluation design did not extend to the IPME team evaluating itself, it is useful here to briefly consider how the IPME assignment worked out in practice. The IPME team developed productive working relationships with COs. **This was undoubtedly helped by the fact that they were deployed not only to quality assure programme monitoring systems but also to provide monitoring-related technical assistance, on a demand-responsive basis.** Informal feedback from UNICEF personnel at all levels indicates that this was much appreciated and COs were generally receptive to advice. Further discussion on the VFM component is provided in Section 3.5.

Regarding the 'accountability' part of the IPME assignment, it was significant that the IPME contract began almost one year later than ASWA itself. It was only at the end of the IPME inception period (in late 2014) that agreement was reached with DFID on the role of national consultants. By the time the appraisals were completed, the programme was already half way through its original implementation period, hence there was limited time within which recommendations to strengthen programme monitoring systems could be implemented. In more than one case, the CO accepted IPME recommendations in principle but did not act on them immediately, indicating instead that they would take them on board for future programmes. Examples include a recommendation to create a common programme database in Bangladesh incorporating data from all three IPs; and to provide more detailed guidance and capacity building support on monitoring for the 13 IPs in Pakistan. In contrast, the Myanmar CO introduced an improved template for recording village and community data following the appraisal, and this was in use by the time of the evaluation, although some other weaknesses in monitoring were yet to be resolved[59].

Overall, the IPME team consider that their engagement with UNICEF has been very productive, and informal feedback from UNICEF received outside of the evaluation tends to confirm this. This said, **there were three limiting factors on the effect of IPME interventions:**

1. A drawback of the demand-responsive approach to TA was that the highest number of support requests came from programmes that were already quite strong, while some others which arguably needed support more, did not request it.

2. Security issues prevented any IPME visits to Yemen, and no field visits could be made in Niger and South Sudan, though some office-based support was provided.

[59] The Myanmar CO later requested IPME support in 2017 to further improve programme monitoring systems.

3. While UNICEF COs accepted in principle the majority of recommendations arising from monitoring systems appraisals and output verification exercises, they were under no real obligation to act on them quickly.

In future, it would be more productive to have IPME (or its equivalent) in place at the programme start to support and quality assure baseline surveys, and to conduct monitoring systems appraisals much earlier, with subsequent rounds of output verification at agreed intervals. In addition, it would help to secure timely action on recommendations if periodic three-way meetings were held between UNICEF, DFID and IPME (perhaps through virtual meetings) to review the latest IPME report(s) and agree on any remedial action to be taken.

Table 29: Guide to Evidence

Evidence Sources	Other: IPME Monitoring Systems Appraisals in 8 countries and synthesis report Ongoing IPME engagement with country offices
Triangulation	ASWA programme globally
Strength of Evidence	Medium: Corroborating evidence from ASWA programme globally

3.8 Other Changes and Innovations (EQ 9,9a)

Table 30: Evaluation Questions

Evaluation Questions Addressed in this Section	
9	What other changes (positive/negative, direct/indirect; intended/unintended) have occurred as a result of ASWA interventions?
9.a	What evidence is there that particular innovations within the ASWA programme are being replicated beyond the initially intended reach of the programme (e.g. outside of geographic areas or target groups)?

3.8.2 Other Changes

3.8.2.1 Whether other changes (positive / negative, direct / indirect; intended / unintended) have occurred as a result of ASWA interventions (EQ 9)

There is insufficient evidence from the three case studies or from analysis across all nine ASWA countries to draw findings and conclusions on what other changes (positive/negative, direct/indirect; intended/unintended) have occurred as a result of ASWA interventions.

Uncorroborated evidence from the Key Informant Interview method is presented for illustration only. In Madagascar, there are anecdotal reports that ASWA school WASH support is leading to increased demand for water by children, inspiring other schools to install their own hand washing points, and leading children to share hygiene practices with their families. In Pakistan, anecdotal reports cite ASWA sanitation support reducing conflict between households because they are not putting waste on each other's land; increasing communities' capacity and mobilisation to access government support on other WASH issues; increasing awareness of disability issues; and increasing school enrolment and attendance. In Myanmar, there are anecdotal reports that ASWA water systems support is improving school attendance through time saved not collecting water. Also, some schools have gained access to safe water through connection to the new piped water system and water tariff funds have been used to hire teachers and repair / improve school buildings. Lastly, in Myanmar access to improved water supplies has enabled expansion / start-up of marketable agricultural produce in some villages.

3.8.3 Innovations

3.8.3.1 Whether particular innovations within the ASWA programme are being replicated beyond the initially intended reach of the programme (EQ 9.a)

There is some evidence from analysis across all nine ASWA countries that **there are a few instances of innovative approaches being used within ASWA (small piped water schemes in Myanmar and Madagascar, use of mobile phones for monitoring in Madagascar and Bangladesh, and**

use of social norms methods to extend of CLTS approaches in Madagascar). However, these are not, so far, being replicated beyond the initially intended reach of the programme.

Table 31: Guide to Evidence

Evidence Sources	Global Analysis (Annex I): EQ 9. a
Triangulation	ASWA programme globally
Strength of Evidence	Medium: Corroborating evidence from ASWA programme globally

3.9 Summary of Lessons Learned

1. The inclusion in programmes like ASWA of countries affected by conflict and/or natural disasters carries the risk that programme implementation will be interrupted; there may also be limited scope for the independent verification of results. These risks need to be factored into programme planning.
2. It is possible to deliver results at scale even in the context of weak government and a dearth of competent NGOs, by deploying contracted project staff directly into government agencies. For these benefits to be sustainable, however, it is vital that the deployed TA does not focus only on direct implementation but also helps to build institutional capacity; and that government in due course appoints sufficient, suitably skilled permanent staff to cover all vital positions.

3. Since UNICEF works via implementing partners and its procurement processes can be quite lengthy, the programme implementation period is often shorter than suggested by programme documents. Whether or not it is explicit in the programme design, an inception period of several months will normally be needed when starting up new programmes.

4. The output indicator for people 'reached' with hygiene promotion was of limited value because it was open to widely differing interpretations and difficult to measure. Reaching people with hygiene promotion is, in any case, an input not an output. In future, it would be more useful to focus on the results of hygiene promotion, which at output level would include the additional number of people gaining access to a hand washing facility near their toilet.

5. It is important to reach consensus at programme start on the level of ambition at outcome level (i.e. for the translation of outputs into outcomes); and on how and when outcome level results will be assessed and reported.

6. In multi-country programmes such as ASWA there is a need for country-specific log frames or results frameworks nested within the global logframe, so that programme focus and the level of ambition at country level are clear. This is especially true for enabling environment outputs since the baseline situation, sector priorities, and intervention types, will vary significantly from one country to the next.

7. The drive to meet ambitious output targets needs to be balanced by attention to securing the funding and other enabling conditions necessary for sustainability.

8. There is consensus within UNICEF globally on the need to monitor both the establishment of enabling conditions for sustainability, and the actual achievement of sustainability. Progress in monitoring sustainability across the nine ASWA countries was patchy, however, and where something was done, the type and scope of activity varied widely. This reflects a lack of consensus on a form and scope of sustainability monitoring that would be both useful and repeatable at agreed intervals, as opposed to one-off studies that are expensive and highly labour-intensive.

9. As with outcomes and sustainability, there is a need for clarity and consensus at the programme design stage on equity objectives and targets, and on how results will be assessed and reported.

10. Consistent VFM budgeting, monitoring, and reporting across country programmes within a funded portfolio presents an additional call on resources and capacity compared to VFM management within discrete programmes. Needs within UNICEF country offices and headquarters to meet this portfolio VFM management challenge need to be assessed and explicitly accounted for within funding proposals.

4. EVALUATION APPROACH

4.1 Introduction

The ASWA programme presented a complicated geographic, organisational, and security context for the evaluation[60]. Therefore, a realistic and practical evaluation approach, balancing what is desirable with what is useful and achievable, was required. In addition, the wider role of the IPME team with ASWA meant that the evaluation could draw on prior knowledge of the programme, but needed to retain a degree of independence[61].

As set out in the evaluation design (see Annex S, Section 5 Approach and 10 Analysis) the evaluation selected a mix methods approach combining case based and theory based elements. The selection was based on: a review of the evaluation purpose and objectives; thorough assessment of the programme attributes and operating context; and our understanding of the range of design options available.

Case Based elements were the primary method pursued through three ASWA country case studies. These were embedded in a larger ASWA programme case study. This allowed the evaluation to:

- Look in depth at specific locations where opportunities for learning are high;
- Explore literal replication[62] by selecting three countries which are predicted to have similar results (as set out in the elaborated ASWA TOC); and
- Look both within a specific case (e.g. at complex factors such as sustainability and enabling environment), across cases, and at the whole ASWA programme (including support from UNICEF HQ and ROs) as the main case to see if this supports the overall TOC for the ASWA Programme.

Theory Based elements allowed the evaluation to map and track the contribution of ASWA to outcomes on the enabling environment for WASH in the TOC at global, regional and country scales. It also allowed the evaluation to map and track ASWA's contribution to behavioural outcomes in terms of HWWS and ODF status in the behavioural change areas of the TOC. Focal areas of interest to DFID in the programme logic of the elaborated ASWA TOC (see Annex B) have also been analysed in relation to the evaluation questions. The findings and conclusions from this analysis are presented in Appendix 1 in Programme Logic tables.

As set out in the evaluation design (see Annex S, Section 7 Country Selection), three of the nine ASWA locations were selected for the country case studies: Madagascar, Myanmar and Pakistan (see Section 5.4.1 below for country sampling strategy).

As part of the country case studies, the evaluation included a post-implementation sustainability assessment in Myanmar (see Annex O) conducted by the IPME team, and drew on the findings from two sustainability checks independently designed and carried out by UNICEF and national governments in Madagascar and Pakistan[63].

In addition to the three country case studies, the evaluation included a comprehensive VFM analysis at the ASWA programme level and country level VFM analyses for all nine countries, as well as the planned

[60] For further information on context see Annex S Evaluation Design Section 2.1, pp. 47-49.

[61] For further information on independence see Annex S Evaluation Design Section 3.1, p. 49.

[62] Literal Replication between cases within a multi-case study can be explored when cases are selected because they are predicted to have similar results (e.g. because in ASWA the case study countries shared a common logframe and TOC).

[63] For further information on the overall approach to sustainability assessment / checks see Evaluation Design Document (Annex S), pp. 112-116.

programme level analysis (see Annex P). This drew on performance and cost data supplied by UNICEF HQ and the end of project data provided by each CO.

Although statistical and participatory methods were not selected to feature prominently within the approach, the evaluation used quantitative methods and participatory working practices to support the mixed qualitative methods[64].

Cross-cutting gender and equity issues have been addressed in the evaluation, firstly through focused evaluation questions that cut across methods (See Table 2 Section 2.4 EQ Nos. 5, 8, 8a, 8b). Gender and wealth disaggregated data were not part of the routine reporting under ASWA, but DFID did request UNICEF to provide this information twice yearly as part of DFID's results monitoring process. Consequently, the evaluation's expectation was that data availability would be variable. Secondly, within the SAM and ROA methods the evaluation took practical steps to ensure equity. Within the SAM beneficiary households and communities were engaged through focus group discussions and structured surveys. Specific focus groups were held with women, men, elderly, and disabled. Within the ROA workshops the inclusion of representatives of beneficiaries who cover dimensions of gender and equity was enabled by making relevant provisions that would facilitate their participation (e.g. security, confidentiality and physical access).

Given that case study-based elements were used, the evaluation incorporated techniques to address threats to external validity so that it could distinguish between country specific and general findings, indicating the extent of external validity and other limitations[65]. This rigorous approach to external validity was anticipated to limit the number of findings that could be shared beyond the ASWA evaluation primary audience (e.g. to other DFID WASH programmes and the wider WASH sector) because many findings were expected to be country specific.

4.2 Process

The design and implementation of the evaluation took place with the collaborative involvement of the DFID WASH Team (including SEQAS review) and UNICEF (including headquarters WASH Team, headquarters Evaluation Office, Regional and the UNICEF Myanmar, Pakistan, and Madagascar COs. The semi-structured interview questions for Key Informant Interviews, the country specific TOCs for Theory of Change Review (TOCR) workshops, and the community level programme activities for ROA workshops were validated by each UNICEF CO before each evaluation mission. The first evaluation mission to Myanmar was also used to test and, as necessary revise, the methods prior to the subsequent evaluation missions.

Other evaluation stakeholders[66] in the case study countries were involved as participants in KIIs, TOCR workshops, and ROA workshops. Selected staff from UNICEF COs in Myanmar, Pakistan, and Madagascar were also interviewed as key informants, participated fully in the TOCR workshops, and were observers at the ROA workshops.

The SAM (see Annex O) took place between February and June 2016, including pilot testing of embedded community and household level data collection tools in March 2016. The Programmatic VFM Assessment (see Annex P) was conducted in February and March 2017. Country evaluation missions, data gathering and preliminary analysis took place between June and December 2016. For the three

[64] For further information on quantitative methods and participatory working practices see Annex S Evaluation Design Section 5.3, pp. 68-69.

[65] For further information on techniques to address threats to external validity see Annex S Evaluation Design Section 5.3, p. 69.

[66] Evaluation Stakeholders in country included government counterparts, implementing partners, service providers, development partners, academics, journalists, rights based organisations, faith based organisations, elected officials, community based organisations, communities and households.

country case studies the evaluation team conducted 67 KIIs, 3 ROA workshops and 4 TOCR workshops that enabled the evaluation to gather views from the diverse range of stakeholders set out in the evaluation design (see Annex Q for a list of consultees by organisation / group). The evaluation team worked freely and without interference from UNICEF, DFID or implementing partners / government counterparts. Table 32 below summarises the activities that took place in the country missions.

Table 32: Summary of ASWA Evaluation Country Mission Activities

Country Mission	Country Mission Team	Scope of Work
Myanmar (20-30 June 2016) • Commercial (Yangon) and administrative (Nay Pyi Taw) capitals • Field visit to Pakkoku and Myaing Townships	• Jeremy Colin (Team Lead) • Carl Jackson (Evaluation Component Lead) • Eain Nyein San (National Consultant) • Myo Myat Thu (National Consultant)	• 24 KIIs in Yangon, Nay Pyi Taw, and a township • 1 ROA Workshop in a Township (41 participants) • 2 TOCR Workshops in Yangon (11 participants) and Nay Pyi Taw (13 participants) • 2 community visits in two villages in two Townships • Presentation of preliminary findings and conclusions to UNICEF CO
Pakistan (8-18 August 2016) • Islamabad • Field visit to Bahawalpur District, Punjab province	• Jeremy Colin (Team Lead) • Carl Jackson (Evaluation Component Lead) • Lara Johnson (Principal Consultant WYG) • Anum Masood (National Consultant) • Ahsan Malik (National Consultant)	• 17 KIIs in Islamabad, Lahore and a district in Punjab • 1 ROA Workshop in a town in Punjab province (18 participants) • 1 TOCR Workshop in Islamabad (14 participants) • 3 community visits in three villages in a District, Punjab Province • Presentation of preliminary findings and conclusions to UNICEF CO
Madagascar (4-15 September 2016) • Antananarivo • Field visit to Analanjirofo region	• Jeremy Colin (Team Lead) • Carl Jackson (Evaluation Component Lead) • Voninala Ranaivo (National Consultant) • Mbola Randrianarivo (National Consultant)	• 26 KIIs in Antananarivo and Mahanoro • 1 ROA Workshop in one commune in Analanjirofo region (26 participants) • 1 TOCR Workshop in Antananarivo (20 participants) • 2 community visits in one commune in Analanjirofo region • Presentation of preliminary findings and conclusions to UNICEF CO

Following the evaluation missions, the Evaluation Co-Principal Investigators (Jeremy Colin and Carl Jackson) and wider IPME team held a triangulation workshop (London, 1-2 November 2016). The objective of this session was to capture IPME Monitoring and Verification and Customised Support (CS) knowledge and triangulate evidence emerging from the evaluation's primary data collection. The workshop assessed the robustness of emerging findings and identified key gaps in evidence requiring additional research. Following the triangulation workshop, further desk analysis of the ASWA programme and the drafting of the country case study reports took place between November 2016 and March 2017.

The evaluation team conducted analysis of evidence collected by each of the methods systematically within nested tables following Yin (2014, Chapter 5) for both triangulation (i.e. confirming and corroborating results reached by one method with other results reached by another method) and/or complementarity (i.e. results obtained by one method helping to better understand those obtained by another method). There were three stages of analysis:

1. **Method-Level:** Evidence gathered by evaluation question for each case study county and at global level for all nine ASWA countries was collated and analysed for each individual method. Findings by question were entered in a pre-structured Method Level table for each method (see

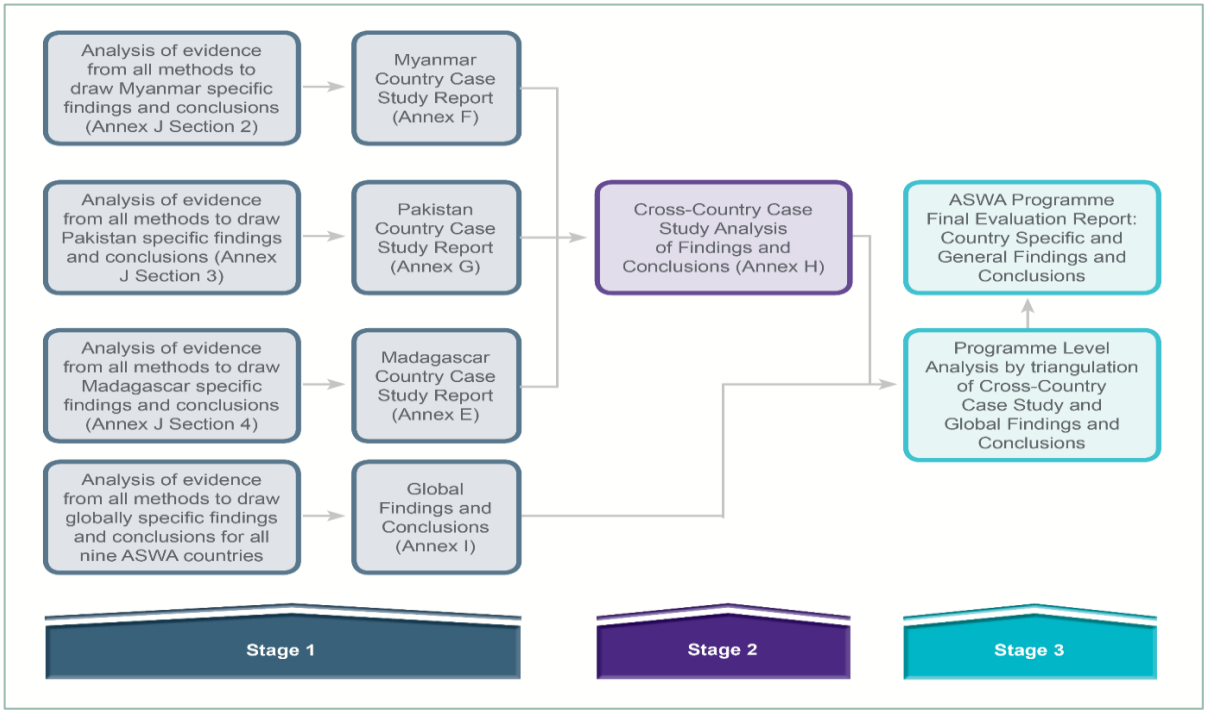
Annex R for example table for KII method). The evaluation team performed analyses by evaluation question across method level findings for each case study country and at global level for all nine ASWA countries. This analysis used triangulation to identify confirmations or corroborations between methods and to identify complementarity to improve understanding of findings from one method with those obtained by another. For each evaluation question the relevant standard for judging performance or descriptive synthesis was applied to enable findings and conclusions to be captured in country level tables (see Annex J Sections 2, 3,4) and a Global level Word table (see Annex I). Separate country case study reports were also produced based on country level analysis to communicate interim findings and conclusions to UNICEF COs (see Annex E – Madagascar, Annex F – Myanmar, and Annex G – Pakistan).

2. **Cross-Country Case Study Level:** Next, the evaluation team performed an analysis by evaluation question across country level findings and conclusions in terms of triangulation. This enabled probing by questions of whether different cases appear to share similar profiles and deserve to be considered literal replications or are contrasting cases. This helped to indicate the degree of robustness of findings and conclusions across all cases and contributed to understanding the external validity of any generalisations that could be made in the third stage of analysis. The resulting cross-country case study findings and conclusions were captured in the Cross-Country Case Word table (see Annex H).

3. **ASWA Programme Level**: Finally, the evaluation team performed an analysis by evaluation question across findings. Firstly, in terms of triangulation to identify confirmations or corroborations between cross-country findings and conclusions (Annex H) and global findings and conclusions (Annex I). Secondly, in terms of complementarity to improve understanding of findings from cross-country level with those obtained from global Level. Empirically observed events were matched to those predicted by the elaborated ASWA TOC to affirm, reject or modify the programme logic. Plausible rival explanations were examined by reference to analysis of contextual data collected. Programme level findings and conclusions are captured in this Final Evaluation Report in Section 3.

These three stages of analysis are represented in Figure 2 below:

Figure 2: Stages of Analysis in the ASWA Evaluation



4.3 Methods and Revisions

Evaluation methods were selected to best respond to five main drivers: the overall evaluation approach (see Section 4.1 above); the evaluation questions and information needed to answer them (see Annex D Evaluation Framework and Section 1.4); achieving a complementary balance between collecting qualitative and quantitative data; a fit with the operational context of the ASWA Programme (see Section 2.2); and the resources and time available for the evaluation provided by DFID.

The evaluation methods were implemented in line with Operational Guidance Notes (see Annex L) in a logical sequence. This ensured consistent application and that evidence and analysis from one method enabled effective implementation of subsequent methods along the evaluation time line.

Three embedded country case studies were the main source of primary data (see Annexes E, F and G). Where primary data could not be collected from stakeholders in the case study countries or from the other six ASWA countries not visited, secondary data sources were identified and reviewed (see Annex K for a full list of documents reviewed by the evaluation). Table 33 below summarises the seven methods selected as per the evaluation design (Annex S), any revisions that occurred during implementation, and the evaluation questions and criteria addressed by each method. The methods are presented sequentially in the order they were used.

Table 33: ASWA Evaluation Methods

Method as per Evaluation Design[67]	Summary of and Rationale for Revision(s)	EQs and Criteria addressed[68]
Country Case Studies (Madagascar, Myanmar, Pakistan)		
Sustainability Assessments (SA) ▪ Post-Implementation SA in Myanmar as part of an in-depth assessment of the quality of the outputs and outcomes delivered under ASWA ▪ Analysis of findings from two Sustainability Checks (SC) independently designed and carried out by UNICEF and national governments in Madagascar and Pakistan ▪ Cross-comparison of these three countries on the service level and a few elements at service authority level	No revisions	2, 2a, 4, 6, 6a-e, 7, 7a-e, 8, 8a-b, 9, 9a, 11, 12 Effectiveness Sustainability Equity Upscaling Context
Rapid Outcome Assessment (ROA) ▪ Specification with UNICEF prior to country missions for each ROA location the ASWA components, actors seeking to influence, change in behaviour or processes sought, behaviours and processes at baseline, and key features of the ASWA Theory of Change that are expected to cause the change ▪ ROA workshops in Madagascar, Myanmar and Pakistan conducted with representatives of beneficiaries ▪ Generates complementary evidence to the behavioural dimensions of the SAs ▪ Collaboratively created detailed visual map of behavioural and process changes in actors with the influences that contributed to and/or blocked them ▪ Collects data on immediate, short-term changes that lead to longer transformative changes as anticipated in a TOC and for capturing the initiative's role in bringing about those results	Specification of key features of the ASWA Theory of Change expected to cause the change was combined with specification of ASWA components as 'Inputs or other actions by UNICEF and its implementing partners that are expected to cause the change'. This was because the TOC itself was insufficiently known / used at community level to be a relevant category for specification in its own right.	2, 2a, 6, 7 Effectiveness Sustainability
Theory of Change Review (TOCR)	No revisions	3a, 4, 7a-e

[67] See pp. 110-126, Section 9 (Methods) of the Evaluation Design Document (Annex S).

[68] As per the ASWA Evaluation Framework, pp. 86-108, of the Evaluation Design Document (Annex S).

▪ Verification of the status of the elaborated TOC with UNICEF prior to country missions ▪ TOCR workshops in Madagascar, Myanmar and Pakistan with the UNICEF Country Office, Implementing Partners, Development Partners and Government ▪ Tests if linking mechanisms and assumptions in the area of Enabling Environment hold true or if other explanations are more plausible		Effectiveness Sustainability
Key Informant Interviews (KIIs) ▪ Conducted in Madagascar, Myanmar and Pakistan with diverse range of informed stakeholders[69] in each country's capital and one operating district ▪ Gathers stakeholder views on a wide range of qualitative evaluation questions	No revisions	1, 1a-b, 2a-b, 3, 3a-b, 4, 4a-4b, 6, 7, 8a-b, 9, 9a, 10, 11, 12 Relevance Effectiveness Sustainability Equity Upscaling Context
All ASWA Countries		
Programmatic VFM Assessment (PVA) ▪ Overall high-level programme VFM analysis of each Country, with restricted possibility to generalise across countries ▪ Assessment of the effectiveness and equity of the overall programme ▪ Qualitative analysis to understand disbursement of funds for each element of ASWA and how this modality overall may impact on VFM ▪ KIIs with UNICEF HQ / ROs ▪ Two VFM case studies from Nepal and Madagascar ▪ Unit cost analysis for Pakistan	Country level VFM analyses for all nine countries conducted. This is because data availability was more comprehensive across countries than anticipated. Did not contribute to answering EQ 7. This is because outcome data was much more restricted than anticipated.	5, 5a-b, 6, 7 Efficiency Sustainability
Global Review of Programme Results (GRPR) ▪ Desk review of available and IPME-verified UNICEF monitoring data (primarily from Annual Reviews and Six-Monthly Reports) ▪ Assesses progress against the ASWA revised logframe output indicator targets ▪ Analysis of Monitoring & Verification Country Reports and a Synthesis Report from the M&V component of IPME ▪ Assesses the extent to which UNICEF country programmes have effective M&E systems that deliver relevant and credible data	Desk review of UNICEF qualitative monitoring information from six monthly country and global reports and Annual Reviews was also conducted. This is because it provided further opportunities to understand contextual factors in ASWA progress.	2, 2a-b, 3a, 4, 4b, 8a-b, 9, 9a, 11, 12 Effectiveness Equity Upscaling Context
Document Review (DR) ▪ Desk review of published reports (sourced from public websites, via email requests to evaluation stakeholders and during evaluation country missions) ▪ Collected documents stored in a structured database and indexed ▪ Information from these documents required for analysis will be compiled into Word or Excel tables with standardised referencing ▪ Qualitative synthesis of information	No revisions	1, 1a-b, 3, 3b, 4, 4b, 8a-b, 10 Relevance Effectiveness Equity Context

In addition to the methods listed in Table 33 the evaluation team drew on knowledge and insights gained during the IPME assignment since its commencement in 2014. Amongst other things, these included: IPME inception visits to ASWA countries and UNICEF headquarters; IPME technical support (including VFM training) to country programmes (both in-country and remote); IPME in-country

[69] For more information on the diversity of stakeholders see Annex Q: List of Consultees.

monitoring systems appraisals and output verification exercises; and IPME participation in two ASWA global meetings.

4.4 Sampling Strategy

The evaluation approach and methods used appropriate sampling strategies in the following areas.

4.4.1 Country Case Study Selection

A two-step selection process led to this choice of the three case study countries. The process reflected the need for operational access to evidence and the likely usability of the evaluation findings. It did not follow a statistical sampling approach because this is not appropriate in case study design (see Yin 2014). The resulting ranking (presented in Table 34 below) selected Pakistan, Myanmar, and Madagascar as the case study countries[70].

Table 34: Criteria Scoring for Country Selection

Country	UNICEF Monitoring Data Availability	Opportunity for Learning	Size of Program	IPME VFM Analysis Depth	Total Score per Country	Ranking
Weighting	**3**	**4**	**1**	**2**		
Bangladesh	4	1	4	5	30	4th
Cambodia	2	4	1	3	29	5th
Madagascar	4	4	4	4	40	2nd
Myanmar	3	5	2	3	40	2nd
Niger	4	3	5	3	35	3rd
Pakistan	5	4	4	3	41	1st

Scoring: 1 Very Low to 5 Very High; Size of Program scale: <£0.5m = 1; <£1m = 2; <£2m = 3; <£3m = 4; >£3m = 5

4.4.2 Myanmar Sustainability Assessment

The SAM method used multistage stratified cluster sampling[71]. The universe was taken to be those villages targeted by the ASWA Programme, for which the total is 225 villages (i.e. clusters)[72]. These villages are considered rural and fall within the four townships of Wetlet, Nwahtogyi, Pauk, and Myaing. These townships represent the first level of stratification. The unit of statistical analysis was the household, and significance is to the township level, to allow comparison of household level indicators between the townships. A sample size was calculated using a 90% confidence and 5% precision, assuming the maximum variance in the primary indicator (p=50%) for each of the four townships. This was increased to account for design effect of 1.5 and a non-response of 10%. This resulted in a final target sample size of 447 households per township for an overall target sample size of 1,788 households[73] across all four townships.

Considering the available resources and the types of interventions included in the ASWA programme in the four townships, the overall number of clusters to be visited was determined to be 12 per township (i.e. 12 per strata) or 48 villages overall. In addition, to ensure a good mixture of interventions in the final sample, the clusters were categorized into three groups (i.e. water supply only interventions, CLTS only interventions, and both water supply and CLTS). The 12 clusters per township were allocated into the three groups in each township considering the proportion that these clusters represented in the total number of clusters in all townships. Proportionate sampling was conducted for the households in each cluster.

[70] For further information on the country case study selection process see Annex S Evaluation Design Section 7.2, pp. 83-85.

[71] For further information on the Myanmar Sustainability Assessment sampling methodology see Annex O (Section 3.3).

[72] There was a total of 41,507 households in these 225 villages.

[73] There was a total population of 48,462 people within the 1,788 households in the sample.

4.4.3 Key Informant Interviews

The KII method is qualitative and so did not use a statistical sampling approach[74]. The evaluation team identified the key individuals within each stakeholder group in each country. The identification process was conducted in advance of country missions by building upon inception phase country visit and country monitoring reports. This identified a group of key informants diverse enough to provide triangulating perspectives on the ASWA activities in each country (e.g. civil society advocates, academics, journalists, private sector as well as UNICEF, government, donors and implementing agents). Efforts were made to secure participation from key informants who cover dimensions of gender and equity. This included making relevant provisions for security, confidentiality and physical access to the interview location. Identification of additional or better key informants, by drawing on information that emerged during individual interviews (a snowballing approach) was also used as practical within the time frame available for each country mission. The evaluation team also added additional key informants identified during the ROA workshop at the beginning of each country mission to explore key changes identified in the ROA visual maps in more depth.

4.4.4 Theory of Change Review

The TOCR method is qualitative and so did not use a statistical sampling strategy[75]. For each country to be visited, the evaluation team identified national organisations directly involved with ASWA, UNICEF WASH, and the wider WASH development and humanitarian sector. The key stakeholder sub-groups sampled were National Government, UNICEF, Implementing Partners, and Development Partners (multilaterals, bilateral, global funds, foundations, NGOs). From a long list the evaluation team selected organisations that were knowledgeable about the national / sub-national enabling environment and diverse enough to provide triangulating perspectives on the ASWA activities in each country. The identification process was conducted in advance of country missions by building upon inception phase country visit and country monitoring reports, and sampling conducted for the Key Informant Interview method. Efforts were made to secure participation from organisations who cover dimensions of gender and equity. This included making relevant provisions for security, confidentiality and physical access to the workshop location.

4.4.5 Rapid Outcome Assessment

The ROA method is qualitative and so did not use a statistical sampling strategy[76]. For each country to be visited, the evaluation team in collaboration with UNICEF Country Office identified organisations in an ASWA implementation area (e.g. province / region) which have a representation, advocacy, or coordination role in relation to intended ASWA beneficiaries (e.g. community based organisations, local councillors, faith based organisations, rights groups – gender, sexuality, disability, age). As well as having such roles, they should be able to both speak on behalf of their constituencies and be knowledgeable about aspects of the wider enabling environment for WASH in that area. From a long list the evaluation team selected organisations that cover a cross-section of intended ASWA beneficiaries. The identification process was conducted in advance of country missions by building upon inception phase country visit and country monitoring reports. Efforts were made to secure participation from representatives of ASWA beneficiaries who cover dimensions of gender and equity. This included making relevant provisions for security, confidentiality and physical access to the workshop location.

4.5 Evaluation Framework

For each evaluation question, the evaluation framework was developed by:

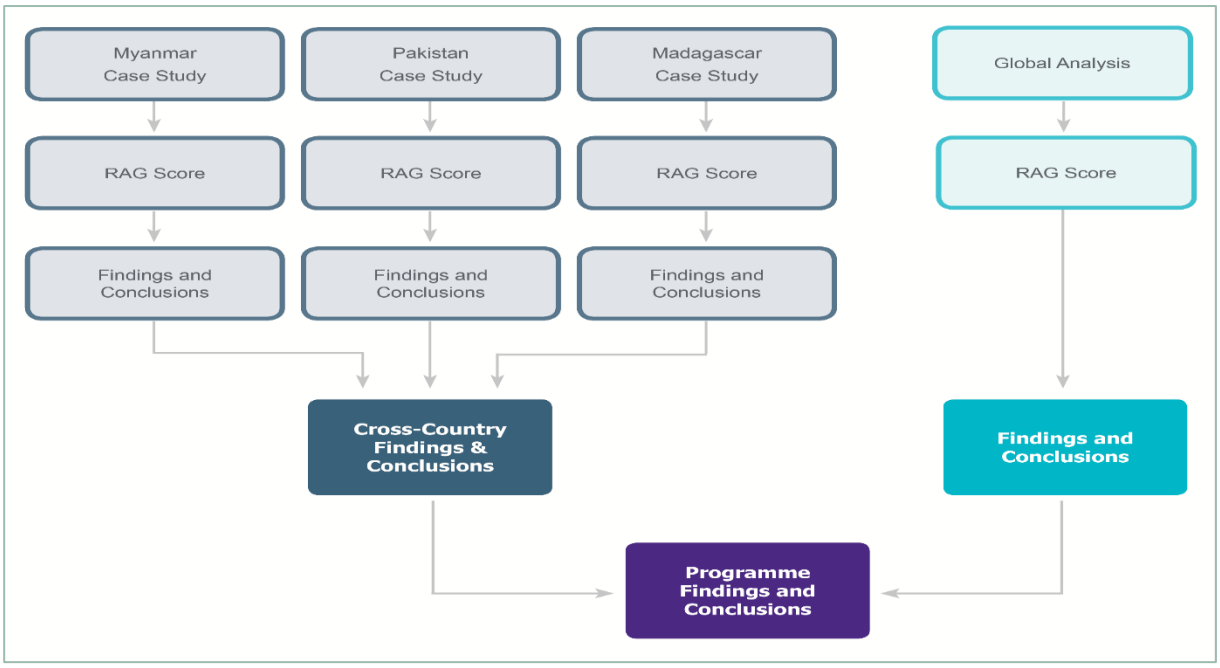
[74] For further information on the sampling strategy used in Key Informant Interviews see Annex L Operational Guidance Notes.
[75] For further information on the sampling strategy used in Theory of Change Reviews see Annex L Operational Guidance Notes.
[76] For further information on the sampling strategy used in Rapid Outcome Assessments see Annex L Operational Guidance Notes.

- Identifying the information that needed to be collected to feed into the analysis that will answer the question, and the source of the data that the evaluation was most likely to have access to (either at the ASWA programme level or in the three embedded country case studies). Two or more sources were identified to support greater robustness through triangulation.
- Selecting the most appropriate combination of elements of case based and theory based methods to collect the information and support analysis.

The ASWA Evaluation Framework (see Annex D) presents for each evaluation question: the information needed by the evaluation to answer each question; the methods to be used; how that information would be used during analysis; and which analytical work streams are used. For questions where an assessment of performance will be made (rather than a descriptive synthesis) the Evaluation Framework also presents the standard to be used (Red, Amber, Green – RAG traffic lights) with criteria for scoring. These were verified with DFID and UNICEF prior to implementation

The Evaluation Team used the 'RAG' scoring for the individual Country Case Studies and for the Global Level Analysis (see Section 10 Analysis in Evaluation Design, Annex S). Figure 3 below sets out the 'RAG' procedure in terms of country, global and programme levels.

Figure 3: RAG Scoring in Analysis Process



In all cases where the RAG traffic lights system is used to judge performance, objectivity was sought by:

- Triangulation of analytical results from two or more methods and three or more ASWA countries.
- Explaining why possible alternative judgements of performance have been rejected.
- Using quantitative definitions of criteria where appropriate that follow the percentage scores for UNICEF Traffic Lights used in the ASWA Annual Review (e.g. Red = <50%; Amber = >50%<75%; Green = >75%).

The application of the RAG traffic lights system for relevant evaluation questions meant that findings and individual conclusions were formulated in a consistent and robust manner. This then facilitated the triangulation of findings and conclusions across country case studies (see Annex H Cross-Country Case Study Analysis, across all nine ASWA countries (see Annex J Analytical Tables, Global Review of Programme Results), and at the Global Level Analysis (see Annex I).

4.6 Limitations

The evaluation design (Annex S) sought to ensure that several potential limitations were pre-empted prior to implementation. Table 35 below summaries potential limitations of the evaluation and how these were addressed by the final evaluation design prior to implementation.

Table 35: ASWA Evaluation Potential Limitations Prior to Implementation

Potential Limitation	How Design Addressed Prior to Implementation
Scope of evaluation in relation to resources and time available.	Original evaluation questions were prioritised in collaboration with DFID, with resources for primary data collection re-focused on primary evaluation questions.
Sufficiency and availability of outcome data from UNICEF and potential over-reliance on key informant opinion.	Limitations of data availability and quality described and discussed in detail and verified with UNICEF prior to implementation.
Evaluation scope / questions not sufficiently aligned with the elaborated ASWA TOC	Priority areas of the TOC were identified by DFID and scope and evaluation questions were focused on these and indicated on the elaborated ASWA TOC
Operationalisation of the RAG traffic lights approach to judging performance potentially subjective / not transparent.	Approach to operationalising the RAG traffic lights approach described in detail and standards for judging performance made more objective in line with UNICEF criteria, but noting limitations to full objectivity.
Adequacy of lengthy (181 pages) Evaluation Design document for communication with evaluation primary audience in DFID / UNICEF.	Evaluation Briefing Document of 14 pages circulated to primary audience prior to implementation to support evaluation utility.
Adequacy of Evaluation Framework in identifying sources of analytical evidence and opportunities for triangulation.	Three Analytical Work streams and how methods support each one described in Evaluation Framework.
Adequacy of Myanmar Sustainability Assessment sampling strategy and approach to analysis.	Sampling strategy and approach to analysis discussed in detail in Myanmar Sustainability Assessment Methodology Appendix to Evaluation Design document.
Adequacy of approach to cross-county comparisons of ASWA performance from national / sector wide Sustainability Assessments in case study countries.	Sustainability Assessment method describes that cross comparison will focus on service level and a few elements at the authority level, but will not attempt sector-level comparison.
Adequacy of Programmatic VFM approach to compiling / analysing data and sufficiency of country level data.	Programmatic VFM method describes analysis is of each country, with limitations expected in possibility to generalise across countries, and that a comprehensive programme wide VFM analysis is not required by DFID. Also, how evaluation team will maintain communication with UNICEF COs and HQ regarding availability and access to VFM data.
Sufficiency of scope of review of secondary documentation beyond DFID and UNICEF sources to understand context.	Global Review of Programme Results and Document Review methods expanded in scope to cover non-DFID / UNICEF secondary sources, including through collection during evaluation missions in case study countries.

Table 36 below summaries further limitations that arose during implementation effecting the data collection methods, evaluation missions and analysis process. It also sets out corresponding mitigation measures the evaluation put in place as part of the adaptive management of the evaluation.

Table 36: ASWA Evaluation Limitations During Implementation

Limitation	Relevant to	Mitigation
Data Collection		
Subjective and fallible recollection on past behaviours and the influences and activities that contributed to and or blocked change. Memory of events that occurred over a period of three years can be inaccurate.	Rapid Outcome Assessment	Triangulation with evidence from Key Informant Interview and Document Review methods.
KIIs have the risk of being un-representative if the stakeholder mapping is incomplete and/or if the individuals do not consent to participate	Key Informant Interviews	Providing sufficient lead-time and resources for mapping; Oversampling to provide alternates when the initial key stakeholder doesn't agree (or isn't available) to be interviewed.
Positive response bias	Key Informant Interviews	IPME made the selection of key informants independently from UNICEF and assured this included a diversity of stakeholders, including those not involved in implementing or benefiting from ASWA.
Evaluation Missions		
In Myanmar, national government stakeholders are located in the administrative capital Nay Pyi Taw, while most development partner stakeholders (including UNICEF) are located in Yangon.	Myanmar Theory of Change Review	TOCR workshops were conducted in both locations by adding an additional day to the duration of the mission.
The number of participants at the ROA workshop was 50% greater than expected. This could have limited the amount of time for individual participants to contribute to the process.	Myanmar Rapid Outcome Assessment	The duration of the ROA workshop was extended on the day from 3.5 to 5 hours.
The duration of the ROA workshop was reduced from three to two hours due to unavoidable logistical constraints (flight cancellation and security curfew). This meant that external actors or factors influencing ASWA in Bahawalpur District could not be fully discussed and the final ROA map could not be verified by participants at the end of the workshop.	Pakistan Rapid Outcome Assessment	Triangulation with evidence from Key Informant Interview and Document Review methods.
During the Theory of Change Workshop, one linking mechanism, that 'local government switches role in enabling sustainable access to WASH from direct investment to promotion or appropriate regulation' was omitted and so its status was not verified by participants.	Madagascar Theory of Change Review	Triangulation with evidence from Key Informant Interview and Document Review methods.
Analysis		
There are constraints on the extent to which programme achievements in the six countries that were not visited for the evaluation can be reviewed. The focus of the IPME assignment prior to the evaluation was on quality assuring programme monitoring systems, validating reported results and providing customised technical support in the areas of VFM assessment, sector monitoring and sustainability monitoring. Beyond these thematic areas, IPME had no contractual remit to review the technical content of the programme or to help ensure that targets were met. Consequently, prior knowledge of ASWA that could be contributed to the evaluation analysis did not cover these areas. Also, security restrictions meant that members of the IPME global team were unable to make field visits outside the capital in South Sudan and Niger, and could not visit Yemen at all. IPME Country Monitors (national	Global Review of Programme Results	Drawing on IPME knowledge and insights gained through direct engagement with the programme during country visits and technical support, both in-country and remote. However, customised support was provided on a demand-responsive basis and consequently we have had significantly more engagement with the programmes in Madagascar, Pakistan and Nepal than with the other six countries, though there have also been opportunities to interact with country office staff via occasional Skype calls and participation in global ASWA review meetings hosted by UNICEF in April 2015 (in Bangkok) and March 2016 (in Antananarivo).

Limitation	Relevant to	Mitigation
consultants) have made field visits in the first two of these countries, but their role was strictly to appraise programme monitoring systems.		Drawing on UNICEF six-monthly reports to DFID produced by COs, along with the UNICEF global reports derived from them. Country report forms not only to provide results data but also make brief comments on programme achievements, highlight issues affecting progress and identify priority actions for the next period.
In the latter half of 2016 ASWA was extended to March 2018, which is beyond the time frame covered by the evaluation. Nevertheless, the timing of the evaluation and cut-off point for data of June 2016 remained unchanged. Therefore, output results and expenditure after June 2016 are not analysed.	Global Review of Programme Results Programmatic Value for Money Assessment	No mitigation. ASWA programme level results to June 2016 for all outputs had already exceeded the Milestone 2 (2015) levels and it was not justifiable to wait for more results that would further exceed these levels. While there may be additional expenditures and further results from individual countries, the overall impact or bias of additional data through December 2016 on analysis is expected to be minimal. Data for most ASWA activity has been captured.
Attributing results of ASWA WASH outputs to ASWA funding alone is difficult and sometimes impossible for the following three reasons. Firstly, in some ASWA countries, DFID is not the only donor to UNICEF WASH programmes and funds from multiple donors are often used for the same or very similar purposes. Secondly, UNICEF contributes its own funds to ASWA interventions and financial systems within UNICEF may not be sufficiently granular to identify exactly where UNICEF and DFID funds are used separately. Thirdly, some ASWA programmes have combined the rehabilitation of past water and sanitation infrastructure with the expenditure of funds for new infrastructure. Without great effort, there is no way to disaggregate lesser expenditures to rehabilitate infrastructure with greater expenditure for new infrastructure. Recognizing the above, DFID agreed with the IPME request to conduct VFM measurements for the entire country WASH programme.	Programmatic Value for Money Assessment	IPME assessed the VFM of entire country WASH programmes for each ASWA country where data was provided. We have drawn evidence-based conclusions at a high level where the comparability of WASH country project data permit. We are transparent about the data sources and limitations, if any.
Although the analysis intended to contribute to answering EQ7 ("What are the major factors and drivers influencing the likely achievement or non-achievement of sustainability of the ASWA objectives for: Sanitation Facilities; Water services; and Hygiene?"), the absence of outcome and impact data for the majority of countries does not allow this. In most countries UNICEF did not plan for or complete studies that generate outcome data.	Programmatic Value for Money Assessment	No mitigation possible.
ASWA beneficiaries are only disaggregated at the programme and country level by gender. The gender disaggregation by country are estimates by UNICEF monitoring and/or partners and reflect country population gender disaggregation.	Programmatic Value for Money Assessment	Retroactive mitigation is not possible. The evaluation notes the weakness of gender disaggregation estimates without verification, and suggests that future programming include specific beneficiary disaggregation by wealth quintile, ethnicity or caste, or other group identifiers that may limit beneficiary access.
RO expenditures for ASWA specific support were not accessible in sufficient detail.	Programmatic Value for Money Assessment	The evaluation team asked UNICEF NYHQ to provide detailed total DFID funding to HQ for ASWA 1; total allocated to HQ, Regions, and to each country; subsequent adjustments made to allocations" and copies of the

Limitation	Relevant to	Mitigation
		three studies conducted with ASWA funding which were commissioned by the RO and HQ level. UNICEF's response was useful, but we were unable to determine in detail how ASWA funds to ROs were allocated or used to support ASWA programmes, except in the most general terms. The three studies funded in part by ASWA were not sent.
The ability to track expenditures against budgets is limited within UNICEF's systems across ASWA programmes. Several key points are identified. Budgeting (Planned Amounts) is not defined at the activity level in VISION; retroactive collection of activity-level financial data requires substantial effort. Where activity financial data are difficult to access, there is less ability for a VFM analysis to understand the process of translating money into outputs. At the activity level, the progress or completion of activities is not monitored as thoroughly and systematically as results at the output level. Results, at the activity and output levels are not reported consistently across country programmes. The level of data disaggregation-financial and results-is insufficient for some VFM analytics at the regional, activity, and beneficiary levels. This data inconsistency undermines the usefulness of comparing the VFM analysis (unit costs, cost-efficiency, regional variances). Quantification of outcome level results, particularly for the enabling environment was hindered by limited quantitative data.	Programmatic Value for Money Assessment	We have recreated original budgets where possible from multiple data sources.
Alignment of methods and indicators across the three sustainability assessments / checks was necessarily limited. Although IPME implemented one (Myanmar) and was involved in another (Madagascar), because two were commissioned separately (Pakistan and Madagascar) by government and UNICEF, IPME could not directly control the design or implementation.	Sustainability Assessment	Alignment of methods and indicators sought through consultation with commissioners and implementers to the extent possible.
The evaluation design envisaged using Microsoft Word Tables to collate and analyse evidence. Word tables could be used for the final presentation of the collated evidence and analysis but were not effective in managing and coordinating the large volume of evidence collected.	Country, Global, and Cross-Country Case Study Analysis	Collation and analysis of evidence was primarily conducted using linked Google Sheets (an online but confidential spreadsheet tool) that was effective in managing and coordinating large volumes of evidence. When finalised these spreadsheets were exported to Word tables for incorporation in the Final Evaluation Report.
Virtually no relevant outcome data (endline versus baseline) was available by June 2016. By the end of 2016, some relevant surveys had been completed, but their usefulness to the evaluation was constrained by a number of factors[77]. As such the evaluation has insufficient evidence to assess outcome performance. Consequently, it is only possible to analyse prospects for sustainable outcomes and draw findings and conclusions in general terms that broadly reflect sector conditions and may not fully reflect the ASWA programme.	Programme Level Analysis	Analysis draws largely from case study country qualitative evidence (from the evaluation methods used during the evaluation missions to the case study countries), supplemented by quantitative and qualitative evidence for those countries from: WASH sustainability assessments / checks for those countries (two of which were sector wide rather than ASWA specific and so only provided a point of reference as they were not necessarily representative of ASWA); a draft post Knowledge Attitudes and Practices

[77] For further information on the lack of outcome data see Section 3.7.2.

Limitation	Relevant to	Mitigation
		(KAP) survey report for one case study country. It also draws on ASWA programme level output target results.

4.8 Inclusion and Research Ethics

All research carried out by WYG and Aguaconsult was conducted in accordance with DFID Ethics Principals for Research and Evaluation (DFID, 2011). It also considered features of the UNEG Code of Conduct for Evaluation in the UN System (UNEG, 2008). All informants consulted consented to participate and were given the option to withdraw at any stage and not to answer questions if they wished. To respect the confidentiality of stakeholders consulted by the evaluation, no individual is named and no precise geographic locations below city level are given in the report or its annexes. Care has been taken to preserve anonymity to the extent possible[78].

The evaluation sought to support the principles of the Paris Declaration on Aid Effectiveness (OECD, 2005 in the following ways. The Sustainability Assessment method sought to support relevance to the WASH sector as a whole rather than seeking to focus on relevance primarily to the ASWA programme evaluation (with some trade-offs in terms of the specificity of data collection instruments and evidence collected). This can be said to have supported country ownership and management of the WASH sustainability agenda and harmonization between DFID and UNICEF. The development of evaluation questions was informed by a combination of the OECD-DAC and UNICEF evaluation criteria, and this can be said to have supported harmonization between DFID and UNICEF in commissioning evaluations. The sourcing of secondary data incorporated that produced by UNICEF and Government to monitor the performance of the ASWA programme, and this can be said to have supported use of local systems. The evaluation team incorporated two national consultants in each evaluation mission who were mentored in the implementation of two contemporary evaluation methods (TOC and ROA), and this can be said to have supported building evaluation capacity within partner countries.

4.9 Communication and Dissemination

Throughout the evaluation, the IPME team maintained regular, bilateral, communication with DFID and the UNICEF COs, Regional Offices and headquarters, to facilitate shared understanding of evaluation plans and actions. An evaluation Joint Reference Group (JRG) between UNICEF and DFID was also proposed, and it has been agreed with DFID that arrangements will be put in place for the JRG following the submission of this report.

The key activities as part of our approach to dissemination of the final evaluation findings are:

1. **End of Evaluation Country Mission Feedback Workshops** – In Myanmar, Pakistan and Madagascar the Evaluation Team held half-day workshops with the UNICEF COs to share provisional findings and conclusions. These aimed to gather feedback, clarify outstanding issues and seek a preliminary validation of findings and conclusions. Verbal feedback from UNICEF on the provisional findings and conclusions was incorporated into further analysis and reporting on the country case studies.

2. **Country Case Study Reports** – Evaluation country missions in Madagascar, Pakistan and Myanmar produced stand-alone reports before their findings and conclusions informed the final evaluation report. These were disseminated to the pertinent COs for comments before the final evaluation report was drafted. The Country Case Studies themselves did not make recommendations, as the evaluation was not designed to provide these country by country, but rather at programme level. Final versions of these reports will be made available to UNICEF COs so that they may share them with stakeholders consulted during the evaluation missions.

[78] For further information on inclusion and research ethics See Annex S Section 12.3, pp. 139-140.

3. **The Evaluation Final Report** – This includes a stand-alone executive summary suitable for online publication and dissemination to stakeholders, research participants and other interested parties via a range channels. Audiences include DFID, UNICEF, Government, Implementing Partners, Development Partners, beneficiary communities, and all of whom will have a keen interest in the findings from key evaluation questions. More broadly, evaluation findings will also be of interest to the global community of practitioners and academics in the WASH sector in the attempt to build the evidence base on what is a sustainable intervention in this field.

Global Learning Workshops – In the overall IPME budget (separately from evaluation budget and workplan) provision is made for two learning events. In April 2015 IPME carried out the first learning workshop in Bangkok. In the Evaluation Design (Annex S), IPME proposed to conduct the second global learning workshop after the evaluation is completed. Devoting a substantial part of this workshop to the Evaluation will aim to gather feedback, clarify outstanding issues and seek validation of findings and recommendations. Our intention is to hold this in late 2017. The modality for this event is still to be agreed with DFID. Options we are considering include:

- Conducting one comprehensive workshop in London with DFID and key UNICEF HQ staff attending in person, while having relevant UNICEF Regional and Country Office staff and stakeholders joining via video conference.

- Holding a series of tailored events with the three UNICEF COs included in the Case Studies (Madagascar, Myanmar and Pakistan).

- Or, holding two regional events, one in Africa and one in Asia for UNICEF and DFID staff.

4.10 Comments on Draft Findings, Recommendations and Lessons

Evaluation stakeholders and end-users have been provided with opportunities to comment on draft findings, recommendations and lessons with actions taken as follows:

- Draft Country Case Study Reports were circulated by email to relevant UNICEF Country Offices (Myanmar, Pakistan, Madagascar) and UNICEF Headquarters prior to incorporation of case study evidence into the triangulation and analysis process. Detailed feedback received from all three country offices and the vast majority of comments and suggested changes, where possible, incorporated into revised Final Country Case Study Reports.

- A Draft Final Evaluation Report was circulated by email to the DFID WASH team and UNICEF Headquarters (accompanied by a request to share it with all UNICEF Country Offices involved in ASWA). Detailed feedback was received from the DFID WASH team and consolidated detailed feedback was received from UNICEF. The vast majority of comments and suggested changes, where possible, were incorporated into a revised second Draft Final Evaluation Report. The DFID WASH team also communicated that thinking from the first draft had already helped to shape the new ASWA programme with UNICEF.

- Feedback on the revised second Draft Final Evaluation Report from DFID's EQUALS evaluation quality assurance process was received by the IPME evaluation team and comments and suggestions prioritised in consultation with DFID WASH team incorporated into a revised Final Evaluation Report.

There were no differences of opinion during the evaluation process within the evaluation team. The team Triangulation Workshop (see Section 4.2 above) provided a framework to surface and clarify opinions, had any differences existed. There was one residual difference of opinion with UNICEF. In relation to the Myanmar Country Case Study Report, the UNICEF Country Office feels that the modification down of the Output 1.1 results (people living in ODF communities) by IPME following on from the Myanmar Sustainability Assessment may not be justified (see Annex F, Section 2.4). The evaluation team remains of the opinion that the modification is justified given that the Myanmar Sustainability Assessment provides robust evidence on ODF and no additional evidence to support maintaining the original output result levels was provided by UNICEF.

5. REFERENCES

This section list documents specifically referred to in the report above. The full list of Documents Reviewed by the Evaluation is included in Annex K.

AAN Associates (2014) 'Evaluation of the UNICEF Sanitation Programme at Scale in Pakistan (SPSP) – Phase 1', Islamabad

Contzen, N., Meili, I. H., et al. (2015) 'Changing Handwashing Behaviour in Southern Ethiopia: A longitudinal study on infrastructural and commitment interventions.' Social Science and Medicine 124: 103-114, Elsevier, Amsterdam

DFID (2011) DFID Ethics Principles for Research and Evaluation, London: DFID. Source: https://www.gov.uk/government/.../dfid-ethics-prcpls-rsrcheval.pdf

DFID (2013a), 'ASWA Business Case'. Source: http://iati.dfid.gov.uk/iati_documents/3985778.odt

DFID (2013b), Terms of Reference for the Process Monitoring and Evaluation Team for the UNICEF programme - Accelerating Sanitation and Water for All in Neglected, Off-Track Countries

e-Pact (2017) 'Evaluation of the WASH Results Programme Mid-Term Evaluation Report', Oxford Policy Management, Oxford

Independent Commission for Aid Impact (2016) Assessing DFID's Results in Water, Sanitation and Hygiene, An Impact Review. Source: http://icai.independent.gov.uk/wp-content/uploads/ICAI-Impact-Review-Assessing-DFIDs-Results-in-Water-Sanitation-and-Hygiene-1.pdf

IPME (2015), Independent Process Monitoring and Evaluation (IPME) Inception Report, March 2015

IPME (2016) Monitoring Systems Appraisals in Niger, South Sudan, Madagascar, Pakistan, Bangladesh, Cambodia and Myanmar, Summary Report, Updated April 2016

Sargsyan et al (2014) Bangladesh WASH sector: large scale impact assessment, Briefing Paper 1983, WEDC, Loughborough University, Loughborough

UNEG (2008) UNEG Code of Conduct for Evaluation in the UN System, June 2008

Yin, Robert Y (2014) Case Study Research: Design and Methods, Sage, London

APPENDIX 1. PROGRAMME LOGIC TABLES

Introduction

Building on the narrative discussion of findings and conclusions in Section 3, where applicable, specific findings and conclusions on the programme logic from related portions of the elaborated ASWA TOC (see Annex B) are presented below. Each table identifies the programme logic elements analysed, findings and conclusions, the extent to which these can be considered analytically generalisable, and the extent to which these can be considered externally valid.

Findings and conclusions are categorised in six ways, as shown in Table A.

Table A: Categorisation of Programme Logic Findings

Affirmed	The predicted programme logic from related portions of the elaborated ASWA TOC is fully supported by the evidence of reality of the ASWA programme.
Partially Affirmed	The predicted programme logic from related portions of the elaborated ASWA TOC is partly supported by the evidence of reality of the ASWA programme.
Should be Modified	The predicted programme logic from related portions of the elaborated ASWA TOC is shown by the evidence of the reality of the ASWA programme across all three country case studies to be inadequate and a revised programme logic may be more appropriate.
Could be Modified	The predicted programme logic from related portions of the elaborated ASWA TOC is shown by the evidence of the reality of the ASWA programme from one or more country case studies and or from analysis across all nine ASWA countries to be inadequate and a revised programme logic may be more appropriate.
Partially Rejected	The predicted programme logic from related portions of the elaborated ASWA TOC is partly challenged by the evidence of reality of the ASWA programme.
Rejected	The predicted programme logic from related portions of the elaborated ASWA TOC is not supported by the evidence of reality of the ASWA programme.

The extent of analytic generalizability is presented as per Table B.

Table B: Extent of Analytic Generalizability

Yes	Indicating that supporting evidence triangulates across all three country case studies, or across one or more country case studies and from analysis across all nine ASWA countries.
Potentially	Indicating that supporting evidence triangulates across two country case studies, but there is insufficient evidence from analysis across all nine ASWA countries to corroborate this.

The option of 'no analytical generalizability' is not included. Triangulation of evidence at prior stages of analysis (see Annex H Cross-Country Case Study Analysis, and Annex I Global Analysis) will have excluded from programme level analysis examples of programme logic where there was insufficient evidence to draw findings and conclusions by further triangulation.

The extent of external validity is presented as per Table C.

Table C: Extent of External Validity

Potentially	Indicating that supporting evidence triangulates across all three country case studies and from analysis across all nine ASWA countries.
Limited	Indicating that supporting evidence triangulates across two or more country case studies and / or from analysis across all nine ASWA countries, but not from both and or sufficiently to corroborate this fully.
Not	Indicating that supporting evidence triangulates across two or more country case studies, but there is insufficient evidence from analysis across all nine ASWA countries to corroborate this.

Programme Logic Tables

Table D: Programme Logic Output 1 (see also Section 3.2.4 above)

Programme Logic	Findings and Conclusions	Analytic Generalizability	External Validity

Increased availability of household technologies and materials (input) – Natural disasters or conflict will not prevent or impair project implementation in the countries or damage sanitation facilities post construction (Assumption A) + Local private sector is able to function / profit points (Assumption N) - Access to basic sustainable sanitation facilities.	Insufficient evidence. Increasing the availability of household technologies and materials for sanitation (i.e. Sanitation Marketing through the private sector) was not a feature of the programme in most countries. Only the South Asian country programmes had sanitation marketing components and some of the related initiatives were introduced under other programmes not covered by the evaluation.	N/A	N/A
Improved access to information for communities and implementers (input) – Access to sustainable sanitation facilities (output).	As above	N/A	N/A

Table E: Programme Logic Outputs 3 and 4 (see also Section 3.2.5 above)

Programme Logic	**Findings**	**Analytic Generalizability**	**External Validity**
WASH in Schools (input) – WASH guidelines for schools represent a reasonable quality of access and costs (Assumption I) – Improved WASH facilities in schools (Output)	Partially affirmed: Some country specific evidence (Pak + Mada) and corroborating evidence from analysis across all nine ASWA countries. Partially affirmed, firstly because UNICEF overachieved the output target for schools with access to WASH by 8-72% (Pak + Mada) and the ASWA output target for schools with access to WASH has been over-achieved in all nine ASWA countries. Despite there being no evidence from outcome indicators for all nine ASWA countries to assess the extent of the ASWA contribution to improved WASH facilities in schools, there is no strong evidence that other WASH programmes constitute a plausible direct rival explanation for these programme level outcomes (see Section 3.4.7.2, EQ10). There is insufficient evidence from the three case studies or from analysis across all nine ASWA countries to assess WASH guidelines for schools represent a reasonable quality of access and costs (Assumption I).	Potentially	Limited

Table F: Programme Logic Output 2 (see also Section 3.2.7 above)

Programme Logic	**Findings**	**Analytic Generalizability**	**External Validity**
Improved access to low cost and innovative water quality and quantity technology (input) – local private sector is able to function / profit points (Assumption N) – Access to sustainable improved water supplies.	Partially affirmed: Some country specific evidence (Pak + Mada) and corroborating evidence from analysis across all nine ASWA countries. Partially affirmed, firstly because UNICEF overachieved the output target for access to improved water supplies by 8-25% (Pak + Mada), and secondly because the ASWA output target for people gaining access to improved water supplies has been over-achieved across all nine ASWA countries. Despite there being no evidence from outcome indicators for all nine ASWA countries to assess the extent of the ASWA contribution to access to improved water supplies, there is no strong evidence that other WASH programmes constitute a plausible direct rival explanation for these programme level outcomes (see Section 3.4.7.2, EQ10). There is insufficient evidence from the three case studies or from analysis across	Potentially	Limited

	all nine ASWA countries to assess local private sector is able to function / profit points (Assumption N). Though some application of innovative technology (solar pumping) and a key role for the local private sector in Madagascar.		

Table G: Programme Logic for Output 5 (see also Section 3.2.8 above)

Programme Logic	Findings	Analytic Generalizability	External Validity
Policies to encourage and promote sustainable WASH improvement (input) – High quality and targeted knowledge and innovation support influences decision makers (Linkage 2) + Coherent WASH Sector systems or service delivery pathways are put in place (Linkage 3) - Increased capacity in public and private institutions to enable sustainable WASH (output)	Partially Affirmed: Strong country specific evidence (Myan + Pak + Mada) but not corroborated by evidence from across all nine ASWA Countries for Linkage 2. Partially Affirmed, firstly because evidence from UNICEF funded technical assistance for studies is shown to have informed government WASH strategy, policies, plans and budgeting (Myan+Pak+Mada). Secondly, because UNICEF enabling environment activities are shown to have supported sector coordination and learning groups (Myan+Pak) and technical advice has been taken up by government (Myan+Mada). Partially Affirmed: Some country specific evidence (Myan + Mada) but not corroborated by evidence from across all nine ASWA countries for Linkage 3. Partially Affirmed because UNICEF support to government is shown to have improved its support for more coherent WASH actor linkages and WASH approaches (Myan+Mada). Insufficient evidence from the three case study countries or from analysis across all nine ASWA countries to assess Increased capacity in public and private institutions to enable sustainable WASH (output).	Potentially Potentially N/A	Not Not N/A
Institutional strengthening: leadership, technical competence, systems, trained staff, sector monitoring, working at scale (input) – Local government switches role in enabling sustainable access to WASH from direct investment to promotion and appropriate regulation (Linkage 6) + Roles and responsibilities at decentralised levels are clarified (Linkage 10) + Comprehensive national M&E systems strengthened in terms of but not limited to reporting on equity of access, implementation to minimum standards, and the use of monitoring data to inform sector planning (Linkage 12) + Local / national capacity for managing inputs, scaling-up delivery of outputs, follow-up, re-triggering and technical support strengthened (Linkage 13) - Increased capacity in public and private institutions to enable sustainable WASH (output)	Insufficient evidence for Linkage 6. This linkage was not a feature of the Myanmar and Pakistan ASWA TOCs. Could be Modified: Some country specific evidence (Pak + Myan) but not corroborated by analysis from across all nine ASWA countries for Linkage 10. Could be modified to account for the fragility of national contexts. This is because UNICEF enabling environment activities have had to take a tactical approach by working with existing unclear roles and responsibilities or not been shown to have significantly improved clarity of roles and responsibilities (Pak + Myan). Partially Affirmed: Some country specific evidence (Myan + Pak) and complimenting evidence from analysis across all nine ASWA countries for Linkage 12. Partially Affirmed, firstly because there is no evidence of UNICEF having strengthened capacity in comprehensive national M&E systems (Myan) and because UNICEF capacity strengthening is shown to have strengthened provincial M&E systems in general and in terms of WASH indicators, but not in terms of equity or at national level. Also, post-ODF monitoring is a gap at local government / community actor level (Pak). Secondly, because UNICEF activities to improve sector monitoring are	N/A Potentially Potentially	N/A Not Not

Programme Logic	Findings	Analytic Generalizability	External Validity
	shown to have taken place in most ASWA countries, but strengthening in terms of reporting on equity of access is shown to have taken place in a few countries only. Partially Affirmed / Could be Modified: Strong country specific evidence (Myan + Pak + Mada) and corroborating evidence from across all nine ASWA countries for Linkage 13. Partially Affirmed, because UNICEF enabling environment activities are shown to have built capacity for WASH ministries and regional departments, including in water systems, CLTS approach and budgeting (Myan+Pak+Mada). Could be Modified to account for the need to analysis and planning of capacity to better address long-term financing (Pak+Mada), systemic overlaps in roles and responsibilities (Pak), and to account for the need for systematic needs assessments and capacity building plans. This is because UNICEF activities to address capacity gaps are shown to have taken place in most ASWA countries and have been welcomed by government, but have often not been based on such assessment or planning.	Potentially	Limited
Support development of community structures to operate and maintain local systems (input) - Increased capacity in public and private institutions to enable sustainable WASH (output)	Insufficient evidence because output targets in this area have not been systematically monitored or reported by UNICEF for all nine ASWA countries	N/A	N/A
Encourage development of viable financing for WASH improvement (input) – Funding leverages additional and sustained public finance to support systems development at all levels (particularly local government) beyond (individual) donor driven project / programme cycles (Linkage 4) - Increased capacity in public and private institutions to enable sustainable WASH (output)	Insufficient evidence because output targets in this area have not been systematically monitored or reported by UNICEF for all nine ASWA countries and because not possible to draw findings and conclusions from case study countries.	N/A	N/A
Develop partnership frameworks for coordination and joint planning (input) – All partners (including NGOs) increase co-ordination and joint learning (Linkage 11) - Increased capacity in public and private institutions to enable sustainable WASH (output)	Partially Affirmed: Strong country specific evidence (Myan + Pak + Mada) and complimenting evidence from analysis across all nine ASWA countries for Linkage 11. Partially Affirmed because UNICEF enabling environment activities are shown to have increased government and development partner coordination and joint learning (Myan+Pak+Mada). This includes advocacy for sector groups and reviews, support to more coordinated and collaborative approach by government, joint learning around methods and approaches). However, in many countries coordination by development partners with government is weak and joint learning is largely informal. Insufficient evidence from the three case study countries or from analysis across all nine ASWA countries to assess Increased capacity in public and private institutions to enable sustainable WASH (output)	Potentially N/A	Not N/A
Increased capacity in public and private institutions to enable sustainable WASH (output) – The ability and willingness of	Insufficient Evidence for Assumption B. Could be Modified: Evidence from analysis across all nine ASWA countries, but not	N/A Potentially	N/A Limited

Programme Logic	Findings	Analytic Generalizability	External Validity
governments and communities to contribute to the programme will not be affected by external economic forces (Assumption B) + Staff turnover in UNICEF WASH teams and their implementing partners during implementation doesn't undermine capacity (Assumption P) – Achievement of all other programme outputs (output)	corroborated by case study analysis. Could be Modified to account for staff turnover in UNICEF WASH teams during implementation affecting continuity. This is because changes and vacancies in UNICEF WASH staff are shown to have affected continuity, but not undermined capacity significantly as other staff stepped up to cover roles.		

Table H: Programme Logic for Output 1 (see also Section. 3.4.2 above)

Programme Logic	Findings	Analytic Generalizability	External Validity
Access to basic sustainable sanitation facilities (output) – Reduced open defecation (Outcome)	Affirmed: Evidence from one country (Pak) and corroborating evidence from analysis across all nine ASWA countries. Affirmed, firstly because increasing latrine access is shown in one country (Pak) to be reducing OD. Secondly, because the ASWA output target for access to sanitation has been overachieved and for people living in ODF communities has been achieved. Despite there being no evidence from outcome indicators for all nine ASWA countries to assess the extent of the ASWA contribution to ODF, there is no strong evidence that other WASH programmes constitute a plausible direct rival explanation for these programme level outcomes (see Section 3.4.7.2, EQ10).	Yes	Limited
Improved community awareness of sanitation (output) – Socio-cultural barriers can be overcome leading to demand creation (Assumption D) + Households improve their behaviour as a result of sanitation and hygiene campaigns (Assumption O) – Communities become ODF (Output) – Reduced open defecation (Outcome)	Affirmed: Strong country specific evidence (Myan+Pak+Mada) and corroborating evidence from analysis across all nine ASWA countries. Affirmed, firstly because use of adapted CLTS methods that address social-cultural barriers are shown to have contributed to progress towards ODF, including increase in availability of latrines and reduction of open defecation (Myan+Pak+Mada). Secondly, because the ASWA output target for people living in ODF communities has been achieved. Despite there being no evidence from outcome indicators for all nine ASWA countries to assess the extent of the ASWA contribution to ODF, there is no strong evidence that other WASH programmes constitute a plausible direct rival explanation for these programme level outcomes (see Section 3.4.7.2, EQ10). Could be modified: Strong country specific evidence (Myan+Pak+Mada). Assumption O that 'Households improve their behaviour as a result of sanitation and hygiene campaigns' could be modified to account for the need for post-ODF follow-up arrangements to support communities to sustain ODF after the sanitation and hygiene campaign phase has ended, and the benefit of social norms methods in achieving improvements in ODF retention. This is because of, ODF slippage risks when such arrangements are not sufficiently in place (in time and resourced) (Myan+Pak), and of use of CLTS extended with social	Yes Yes	Limited Limited

	norms methods being shown to have led to an improvement in ODF retention (Mada).		

Table I: Programme Logic for Output 3 (see also Section. 3.4.3 above)

Programme Logic	Findings	Analytic Generalizability	External Validity
Improved community awareness of hygiene practices HWWS (output) – Socio-cultural barriers can be overcome leading to demand creation (Assumption D) + Households improve their behaviour as a result of hygiene campaigns (Assumption O) + Soap or ash is used in handwashing (Linkage 7) – Improved hygiene practices (outcome)	Affirmed: Strong country specific evidence (Pak+Mada+Myan) and corroborating evidence from analysis across all nine ASWA countries. Affirmed, firstly because the number of households that have soap or other cleansing agents is shown to have increased (Pak+Mada) or has been confirmed (Myan). Secondly because the ASWA output targets for people reached by hygiene education programmes and for people with soap / ash available by their toilet have both been over-achieved. Despite there being no evidence from outcome indicators for all nine ASWA countries to assess the extent of the ASWA contribution to adoption of HWWS, there is no strong evidence that other WASH programmes constitute a plausible direct rival explanation for these programme level outcomes (see Section 3.4.7.2, EQ10). There is insufficient evidence from the three case studies or from analysis across all nine ASWA countries to assess Socio-cultural barriers can be overcome leading to demand creation (Assumption D) or Households improve their behaviour as a result of hygiene campaigns (Assumption O).	Yes	Limited
Water is used in handwashing (Linkage 8) – Improved Hygiene Practices (Outcome)	Affirmed: Some country specific evidence (Myan+Mada) and corroborating evidence from analysis across all nine ASWA countries. Affirmed, firstly because households are shown to have handwashing facilities with water (Myan+Mada). Secondly because the ASWA output target for people with soap / ash available by their toilet has been over-achieved. Despite there being no evidence from outcome indicators for all nine ASWA countries to assess the extent of the ASWA contribution to adoption of HWWS, there is no strong evidence that other WASH programmes constitute a plausible direct rival explanation for these programme level outcomes (see Section 3.4.7.2, EQ10).	Potentially	Limited

Table J: Programme Logic for Output 4 (see also Section. 3.4.4 above)

Programme Logic	Findings	Analytic Generalizability	External Validity
Improved WASH facilities in schools (output) – Reduced open defecation (outcome)	Partially affirmed: Evidence from analysis across all nine ASWA countries, but insufficient evidence from the three case studies to corroborate this. Partially affirmed because the ASWA output target for schools with access to WASH has been over-achieved and for people living in ODF communities has been achieved. Despite there being no evidence from outcome indicators for all nine ASWA countries to assess the extent of the ASWA contribution to ODF, there is no strong evidence that other WASH programmes constitute a plausible direct rival explanation for these	Potentially	Limited

	programme level outcomes (see Section 3.4.7.2, EQ10).		

Table K: Programme Logic for Output 2 (see also Section. 3.4.5 above)

Programme Logic	Findings	Analytic Generalizability	External Validity
Water points remain functional (output) – Rural populations in the target countries are not displaced as a result of natural, political or economic forces (Assumption E) - Increased consumption of safe water (outcome)	Could be modified: Some country specific evidence (Myan+Mada). Could be modified to account for choices not directly related to safety affecting increased consumption of safe water. This is because choice is a factor in continued use of unsafe sources some of the time even when an improved source is available. The choice to use unsafe sources can be for reasons of taste (Myan) or to only use safe sources at critical times (Mada). There is insufficient evidence from analysis across all nine ASWA countries to corroborate the above because baselines and special studies to assess ASWA outcome indicators have not been planned and / or completed by UNICEF for all nine countries. Partially affirmed: Evidence from analysis of all nine ASWA countries. Assumption E that 'Rural populations in the target countries are not displaced as a result of natural, political or economic forces' is partially affirmed. This is because people in ASWA locations are not reported to have been displaced despite insecurity, conflict, political instability, elections, natural disasters, and epidemics having been present in many countries. There is insufficient evidence from the three case studies to corroborate this finding.	Potentially Potentially	Not Limited

Table L: Programme Logic for Equity Targeting (see also Section. 3.6.2 above)

Programme Logic	Findings	Analytic Generalizability	External Validity
Increased capacity in public and private institutions to enable sustainable WASH (output) - Better targeting of sector resources to provide WASH services to the poor, marginalised and vulnerable populations can be achieved (Assumption J) + Unserved / underserved populations can be prioritised to gain access to WASH services (Assumption L) - Achievement of all other programme outputs (output)	Affirmed: Strong country specific evidence (Myan+Pak+Mada) and corroborating evidence from analysis across all nine ASWA countries. Affirmed, firstly because UNICEF geographic targeting of Townships / Union Councils / Regions is shown to have selected some of the poorest Townships / Union Councils / Regions as ASWA locations (Myan+Pak+Mada). Secondly, because the ASWA output targets for sanitation, water, hygiene, and school WASH have been achieved. Despite there being no evidence from outcome indicators for all nine ASWA countries to assess the extent of the ASWA contribution to increased capacity in public and private institutions, there is no strong evidence that other WASH programmes constitute a plausible direct rival explanation for these programme level outcomes (see Section 3.4.7.2, EQ10). Should be modified: Strong complimenting country specific evidence (Myan+Pak+Mada). Assumption J that 'Better targeting of sector resources to provide WASH services to the poor, marginalised and vulnerable populations	Yes Yes	Limited Not

	can be achieved' should be modified to account for the need to robustly identify poor, marginalised and vulnerable populations, monitor the outcome of targeting on them, and specify women and girls as a specific group. This is because identification was partial, monitoring did not take place and women and girls were not targeted as a specific group (Myan+Pak+Mada). Simply referring to 'better targeting' without defining what this entails does not adequately make technical expectations explicit.		

Table M: Programme Logic for Equity Promotion (see also Section. 3.6.3 above)

Programme Logic	**Findings**	**Analytic Generalizability**	**External Validity**
Increased capacity in public and private institutions to enable sustainable WASH (output) - Better targeting of sector resources to provide WASH services to the poor, marginalised and vulnerable populations can be achieved (Assumption J) + Results can be achieved in greater equity of access to WASH services (Assumption K) + Unserved / underserved populations can be prioritised to gain access to WASH services (Assumption L) - Achievement of all other programme outputs (output)	Should be modified: Strong country specific evidence (Myan+Pak+Mada) and corroborating evidence from analysis across all nine ASWA countries. Assumption J that 'Better targeting of sector resources to provide WASH services to the poor, marginalised and vulnerable populations can be achieved' could be modified to account for the need to build capacity on existing guidance / plans and good practice / insights in equity identification, targeting and monitoring (including gender) relevant to WASH. This is because, firstly such guidance / plans and good practice / insights in UNICEF are shown not to have been sufficiently drawn upon (Myan+Pak+Mada). Secondly, because UNICEF HQ / RO support to countries is not shown to have fully led to process and tools being in place to identify, target and monitor equity.	Yes	Potentially