Common WASH 4W matrices and beneficiary double counting

A brief paper about current limitations and improvement suggestions

Introduction:

Progress against targets in HRPs are often based on numbers from 4W matrices. This tool in its current state is however inadequate for this purpose. There are basic logical reasoning flaws in the way beneficiaries are counted that systematically result in significant double counting and cluster target over achievements across responses. For example, it has been seen that the cluster has reached more than the entire population of a country in some cases or 500% of its targets. What usually happens at that point, is that the person in charge dismisses these numbers and establishes new ones based on difficult approximations and with no solid methodology. The numbers that are then presented to donors, other agencies, OCHA and partners consequently have very important margins of error and their reliability and usefulness become severely questionable.

Problem Statement:

The 3 main issues causing double counting are temporal, sectoral and geographical.

- **Temporal issue**: WASH response activities have different time spans. This dimension is however never taken into account when summing up beneficiaries in the "Number of reached beneficiaries" columns. Example: a borehole produces an amount of water every day that reaches a certain number of beneficiaries over a long period of time (let's say 500 people a day for a month). The line describing this activity is inserted once into the matrix. Bottled water distribution reaches another amount of beneficiaries over a limited period of time (let's say 500 bottles are distributed that benefit 500 people). That is also inserted in the matrix with 500 reached beneficiaries. But as that distributions is repeated, the activity is therefore repeated in the matrix and lines are added. If that activity is repeated every day for one week in the same location, it adds up to 500*7 = 3'500 people in this example. But it really only is the same 500 people.
- **Sectoral issue**: WASH is (at least) 3 domains in one unlike some other humanitarian sectors. Even though water, sanitation and hygiene are complementary and interlinked, they are still separate sub-sectors. Some beneficiaries from water activities will be the same as beneficiaries from hygiene and/or sanitation activities. Different lines in the matrix will address the same beneficiaries causing double counting.
- **Geographical issue**: partner A indicates in the matrix soap distributing in one location for example. Partner B indicates building hand washing stands in the same location for example. These activities from the same hygiene WASH subsector will be in two different line in the matrix and indicated beneficiaries by the two partners will be added up in the current ways of calculating but they might be the same beneficiaries.

In addition, relying on unverifiable partner beneficiary estimates, relying on different and un harmonized methodologies partners use for reporting, trying to capture too much information into one tool (future, present and past projects, beneficiaries targeted and reached, implementing partners and donors, etc.) also complicates the process.

Potential Solutions:

While double counting cannot be completely eliminated due to the complex nature of humanitarian work, it can be minimized with more pragmatic approaches and by using currently available technology.

Beneficiary counting should be entirely removed from the 4Ws to cut out unreliable estimates at the source and calculations should be performed separately by taking into account the amount of services stated in the 4Ws (number of bottles, aquatabs, hygiene kits distributed, cubic meters of water produced, etc.) and comparing those with current norms and standards (Sphere for example).

Implementing partners, donors and planned projects should be part of separate tools perhaps included in capacity mapping or budgeting exercises. 4Ws should be focused only on the response and contain only activities that have been implemented.

- **Solution for temporal issue**: change the unit of beneficiaries. On the same model as [man*hours] are used to calculate necessary resources to accomplish a project [beneficiary*days] can be used in the aid world.
  Random Examples:

  In a given sudden onset crisis, let's say the flash appeal calls for 250'000 people in need of assistance over the next 2 months and the sector targets all the people in need. So that would be 250'000 [people]*2[months]*30[days per month] = 15'000'000 [beneficiary*days].

  – 1 water treatment unit is installed that produces 20'000 [liters] of water per day and runs for 2 months with a minimum sphere standard of 7.5 liters per person per day:

  $$20'000\left[\frac{l}{d}\right] \times 60[d] \div 7.5\left[\frac{l}{b*d}\right] = 160'000[b*d]$$

  – 5000 aquatabs 33 mg that clean 5 liters each are distributed every day but only for 1 months:

  $$5000\left[\frac{tabs}{d}\right] \times 5\left[\frac{l}{tabs}\right] \times 30[d] \div 7.5\left[\frac{l}{b*d}\right] = 100'000\,[b*d]$$

  – 500 people are sensitized to good hygiene practices. Assumption is that they will remember their sensitizing for the whole duration of the flash appeal.

  $$500[b] \times 60[d] = 30'000[b*d]$$

  – 500 bars of 250 grams of soap are distributed in 1 day with a sphere standard of 250 grams per person per month:

  $$500\left[\frac{bars}{d}\right] \times 250\left[\frac{g}{bars}\right] \times 1[d] \div 250\left[\frac{g}{b*m}\right] \times 30\left[\frac{d}{m}\right] = 15'000\,[b*d]$$

  – 50 latrines are installed. Assumption is that they will stay functional during the whole duration of the flash appeal. Sphere standard of 20 people per latrine.

  $$50[lat] \times 60[d] \times 20\left[\frac{b}{lat}\right] = 60'000\,[b*d]$$

  $b = beneficiairies$ $\quad$ $m = month$
  $d = days$ $\quad$ $lat = latrines$
  $l = liters$
  $g = grams$

And such an equation is built for every activity, either with a sphere standard and/or with an assumption.
Then all [beneficiary*days] of the different activities are summed but by subsector (see below) and calculated towards the total number of [beneficiary*days] from the flash appeal producing a percentage of achievement.
In the above example, the 5 activities have reached 2.4% of the total beneficiaries or ~6'000 beneficiaries if the 2.4% is reapplied towards the 250'000 targeted people or if the total number of [beneficiaries*days] is divided by the time span of the flash appeal (60 day). With the current method, without taking in the temporal dimension and by applying the same standards and assumptions, it would be ~8´000.

- **Solution for sectoral issue**: systematically calculate number of reached beneficiaries (according to above methodology) but per subsector exclusively. This is already the case in most settings as there are different strategic objectives per subsector in the HRP with subsequent indicators. But a larger issue here is that general people in need from HNOs/HRPs are not separated by WASH subsectors. In the flash appeal example from above with 250'000 targeted people, there is no break down by subsector which is the case in most HNOS/HRPs. Is the sector targeting 250'000 for water, 250'000 for hygiene and 250'000 for sanitation? Is that 750'000 people targeted? Or is it for example 100'000 for water, 100'000 for hygiene and 50'000 for sanitation? In the first stages of planning, the difference between subsectors must be clear to then allow calculation of progress against targets. In current practices, there is this logical reasoning flaw from the very start that then induces significant double counting. To target beneficiaries as a whole for the WASH sector does not make any mathematical sense if numbers are not broken down by subsectors
  One common total target number for the whole of WASH can still be presented in the HNO/HRPs but the breakdown by subsectors must always be explained in parallel.

- **Solution for geographical issue**: this issue proves to be very complex and cannot be entirely addressed with a simple 4W matrix. It is however possible to integrate assumptions in calculations to reduce double counting. A decision tree can be built to choose what type of calculations should be performed depending on the level of geographical precision the data is reported into the matrix. Whether it is at a macro or micro level. For example:
  - A partner reports doing an activity on an admin 1 or admin 2 level (department, commune, payam, kebele, governorate, district, etc.), which is quite broad and macro. Another partner reports the exact same activity in the same macro location. Assumption is that there is enough place so that partners do not overlap and beneficiaries are added.
  - A partner reports doing an activity in a specific location (camp setting or small locality), which is quite micro. If another partner reports doing an activity in the same WASH subsector in the same location at the same period of time, then maximum of beneficiaries of the two lines are taken into account (and not the sum).

In resume, if the time component is included in the matrix, if calculations are systematically disaggregated by subsector and if assumptions based on macro or micro level of information are taken into account, I believe that double counting can be significantly reduced and that numbers presented to donors, other agencies, OCHA and partners will be more reliable and will enable better gap analysis, coordination and over all response.

Flash appeals, HNOs and HRPS can still be produced the same way, it is still possible to provide a general unique number of targeted people and reached beneficiaries for all of WASH but those numbers need to be detailed by subsector and the way we gather data and calculate progress internally on the cluster side has to be improved. I have myself never seen any methodology explaining how to calculate the progress, I hope this paper can be a beginning. Tools need to be amended and a workflow explaining the methodology needs to be established and rolled out to all countries if we wish to improve within the sector.

General steps for eventual implementation:

1. A simplified Excel 4W template spreadsheet (or google sheet or kobo or any other online data entry system – depending on contexts) is produced (see example below). The architecture can be easily created but the underlying activities in the drop down menus, the units tied to the activities, and the indicators tied to activities will need to be thoroughly examined to be able to create a master template applicable to most contexts.

WHO?	WHERE?	WHEN?	WHAT?					COMMENTS
Organisation	Admin Units	Date of completic	WASH Domain	Activity	Unit (automatic fill)	Quantity	Number of times (or days) activity has been performed	Remarks
			Water	Water trucking	[m3/day]	5,000	30	
			Water	Water bottle distribution 1 liters	Number	1,000	5	
			Water	Aquatab 33 mg distribution	Number	500	5	
			Hygiene	Soap 250 gr distribution	Number	500	1	
			Hygiene	Hygiene promotion	Number of people	500	1	
			Sanitation	Emergency latrine installation	Number	50	1	

2. A backend for data analysis is created with a real database management system (e.g., not Excel) enabling SQL queries and advanced scripted calculations.
   Access for example (a large portion of computers that have Word, Excel and outlook also have Access installed) or any free database management systems. Formulas taking in account norms and standards and assumptions are pre-established in the database. Simplified user friendly panels (or masks) are also created within the database system to enable modification of standards and assumptions. Simplified user friendly panels or processes are also created to enable automatic subsequent analysis of data.
3. A data visualisation system is also preconfigured to be able to integrate and present the result of the analyses from the database. This could be done with Publisher (like Access, installed on a large portion of computers that have Word and Excel) or with Tableau (free or paying version) or Illustrator (no free version).

Once this system is developed, the following steps can be taken by in country IMOs:

4. The 4W template is sent to partners for data collection. Less data is asked therefore simplifying data entry and reporting but more precise data is asked regarding quantities. This could be a limitation as partners will need to quantify their activities precisely. But nothing changes regarding the current status as this is already a challenge today regarding beneficiaries inputs. It will be the role of the in country IMO to obtain and refine that data as it already is today.
5. All collected data through the 4W is imported into the preconfigured database management system.
6. IMOs specify which assumptions and standards are in place for the given context with the user friendly panel included in the database management system.
7. IMOs perform automated calculations through the user friendly panels.
8. Results are pushed to the visualization system.

To review the whole system as suggested in this paper represents an important amount of work. I estimate it would take an experienced IMO with advanced skills about 3 month's full time. It could be done through a consultancy.
The terms of reference of the consultancy would include studying what technology/ies is/are the most adequate, creating the different systems and establishing a workflow with step by step instructions.
Such a system could then be included in the Information Management Toolkit and pushed out to different country platforms.
Ideally, it would be a 6 months consultancy, that would include the creation of the system, a pilot phase, a tool adaptation phase and then in country training of select IMOs (in all L3s for example).

Such a system could then be managed by IMOs with basic skills that would follow step by step instructions with pre designed templates all along the way. The IMO would go into the information management toolkit and download the package then follow the workflow. No coding or SQL querying or anything complex required (all pre built in). The IMO would just follow instructions, integrate admin boundaries from HDX for his/her country, verify or adapt assumptions and standards and then start data collection. After data cleaning and compilation, the IMO would run analyses with the pre included equations based on chosen context standards and assumptions.

A few important caveats however regarding this suggested system:

- Context specific modification (for example, including new activities in the drop down menus of the 4W that haven't been pre-established and therefore new equations, etc.) can only be done by an IMO with advanced skills (adaptation of the 4W to different administrative boundaries for different countries is however simple). It is nevertheless also possible to discard this difficulty by deciding that partners only report on pre-included activities. If all currently known most important WASH activities across the world are included (by building on the indicator registry for example and GWC IMO experience), auxiliary and smaller activities that don't usually represent a lot of beneficiaries could simply be neglected.
- This system does not allow for SADD of beneficiaries (but the current system does not really either, as zero current 4Ws in the world contain complete and reliable SADD data). SADD could however be extrapolated through other assumptions by statistical means.
- Translation of the system into other languages than the one it would be built in could take an additional month of work per additional language.

As 4Ws remain a central piece of any information management systems in humanitarian responses, I believe investing sufficient time and resources in improving the way these matrixes function would be of an important benefit.

A more consistent approach based on current standards and not on partner beneficiary estimations would help IMOs deliver more pragmatic and evidence based numbers. These numbers could then be in turn compared to general population data, to other existing infrastructure data (water system networks, waste water networks, etc.) to eventually determine coverage gaps in a more coherent way.