

Use of Outcome Harvesting for monitoring in Dialogue and Dissent alliances: findings from a survey and discussions

Report author: Richard Smith | RDS Consulting Ltd | rdsmith27@gmail.com

Commissioned in 2020 by:



About the author

rdsmith27@gmail.com | Profile: <http://uk.linkedin.com/pub/richard-smith/32/369/879>

Dr Richard Smith originally trained as a savanna ecologist before moving into programme management in international environment and development organisations. He started using Outcome Mapping (OM) in 2010 for strategic planning and later that year commissioned an Outcome Harvesting evaluation of an advocacy and capacity development network. Through this evaluation, he gained first-hand experience of the demands Outcome Harvesting (OH) puts on programme staff as well as the highly valuable insights and evidence the tool can bring to light. Inspired by these experiences and ongoing research and practice, Richard has, since 2011 worked as an independent consultant directly supporting some 25 organisations, programmes and funding programmes with outcome-oriented design, monitoring, evaluation and learning. He specialises in the practical application of OM concepts, in particular through OH. He enjoys sharing his OH experience and insights by providing training and coaching. In addition, he participates in the Outcome Harvesting Community Forum and through his role as a steward and board member of the Outcome Mapping Learning Community, he shares responsibility for the ongoing development of this learning platform, including its rich resources on OH.

Table of Contents

1	INTRODUCTION	1
1.1	BACKGROUND	1
1.2	THE USE OF OH BY D&D STRATEGIC PARTNERSHIP ALLIANCES	1
1.3	SCOPE AND STRUCTURE OF THIS REPORT	2
1.4	HOW TO USE THE REPORT	2
1.5	SURVEY RESPONDENTS	2
2	MONITORING CHOICES.....	4
2.1	EXPERIENCE OF USING OH FOR MONITORING.....	4
2.2	TRAINING	5
2.3	PURPOSE(S) OF USING OH FOR MONITORING	5
2.4	THE MAIN REASONS FOR CHOOSING OH	6
2.5	WHY OTHER APPROACHES WERE ALSO NEEDED	6
2.6	USE OF THE SIX OH STEPS	6
3	HARVESTING OUTCOMES	8
3.1	FREQUENCY OF HARVESTING.....	8
3.2	HOW OUTCOMES WERE HARVESTED.....	9
3.3	LESSONS LEARNED	10
3.3.1	<i>Practices to maintain</i>	10
3.3.2	<i>Practices to change</i>	11
4	REVIEWING OUTCOMES	14
4.1	HOW OUTCOMES HAVE BEEN REVIEWED	14
4.2	QUALITY AND COMPLETENESS OF FINAL OUTCOME STATEMENTS	15
4.3	LESSONS LEARNED	17
4.3.1	<i>Practices to maintain</i>	17
4.3.2	<i>Practice to change</i>	17
5	SUBSTANTIATION	19
5.1	HOW OUTCOMES WERE SUBSTANTIATED.....	22
5.2	LESSONS LEARNED	23
5.2.1	<i>Practices to maintain</i>	24
5.2.2	<i>Practices to change</i>	24
6	ANALYSIS AND USE OF HARVESTED OUTCOMES.....	25
6.1	WHO PARTICIPATED IN ANALYSIS OF OUTCOMES.....	25
6.2	USE OF SOFTWARE IN ANALYSIS.....	26
6.3	USE OF HARVESTED OUTCOMES.....	26
6.4	LESSONS LEARNED	27
6.4.1	<i>Practices to maintain</i>	27
6.4.2	<i>Practices to change</i>	27
7	ASSESSMENT OF OH EXPERIENCE	29
7.1	ADVANTAGES OR WHY USE WILL BE CONTINUED	29
7.2	DISADVANTAGES AND DIFFICULTIES OR WHY USE WON'T BE CONTINUED.....	30
7.3	EXPERIENCES USING OH IN DIFFERENT CONTEXTS	32
8	LEARNING EVENT DISCUSSIONS.....	33
8.1	DISCUSSION 1: OH DATA QUALITY.....	33
8.2	DISCUSSION 2: FROM USE FOR REPORTING TO USE FOR LEARNING	39
8.3	DISCUSSION 3: SUSTAINING OH IN ORGANISATIONS	40

ANNEX A: SURVEY RESPONDENTS	42
ANNEX B: THEMES FOR EXCHANGE AND LEARNING	44

Acronyms

D&D	Dialogue and Dissent Strategic Partnership
HQ	Headquarters
M&E	Monitoring and Evaluation
MFA	Dutch Ministry of Foreign Affairs
OH	Outcome Harvesting
OM	Outcome Mapping
MSC	Most Significant Change
(P)MEL	(Planning), Monitoring, Evaluation and Learning
SMART	The OH adaptation of Specific, Measurable, Achieved, Relevant, Timely
ToC	Theory of Change

1 Introduction

1.1 Background

Between 2016 and 2020, 25 Dutch coalitions/alliances of civil society organizations have implemented programmes under the Dialogue and Dissent (D&D) Strategic Partnership framework of the Dutch Ministry of Foreign Affairs (MFA). Early in 2020, some 21 of the 25 alliances, plus RNW-Media, decided to compare their experiences with Outcome Harvesting (OH) with the intention of learning for improvement of their future monitoring practices. This report presents the findings of a survey developed by a voluntary OH task force of some D&D coalition members (specifically, M&E staff of Cordaid, Hivos, NIMD, Oxfam Novib, PAX, SNV, Wetlands International and WWF) with support from the independent consultant who authored this report, Richard Smith. The author has added his own observations to the findings, drawing on his experience to indicate good practices and potential discussion points. Considering that much use of OH has been in the context of evaluation, the monitoring experiences shared by the 22 respondents in this report form a potentially valuable learning resource for the organisations involved and more widely.

1.2 The use of OH by D&D Strategic Partnership alliances

The D&D framework provided the 25 alliances with resources for lobbying and advocacy. The great majority of these alliances adopted Outcome Harvesting (OH) as a monitoring approach, some following OH more strictly, others less so¹. The use of OH was a natural choice because of the complexity inherent in lobbying and advocacy interventions: there is considerable uncertainty about which activities of which actors will achieve change as well as uncertainty about what precise changes can be achieved through the influence of lobbying and advocacy. Four characteristics of OH make it particularly well suited for monitoring lobbying and advocacy:

1. **OH is 'actor centred'**. This means that OH seeks to measure changes at the level of institutions or individuals. In other words, it seeks to measure changes among the targets of lobbying and advocacy.
2. **Outcomes as behaviour change**. OH systematically documents outcomes defined as changes in behaviour influenced by an intervention. Lobbying and advocacy that is successful influences the behaviour of actors.
3. **Influence and contribution**. OH assumes that an intervention is but one influence on the behaviour of the actors whose behaviour changes it seeks to describe. In lobbying and advocacy, it is normal that there are other civil society, private sector or state actors over and above those in your intervention who may also contribute to the behaviour changes you seek. In OH, it is essential that plausible influence of an intervention is demonstrated but not that this is the only influence leading to an outcome.
4. **Start from what has changed**. An outcome harvest seeks to describe outcomes that have actually materialised regardless of whether they were planned or not. In this way, OH can capture unexpected outcomes, something that is common in lobbying and advocacy. OH can also be useful for describing, and learning from, the smaller, initial behaviour changes that are often necessary before a final e.g. policy level change is realised.

A fifth characteristic of OH – **that it is a participatory M&E approach** - was, as we shall see in the survey results, widely appreciated because of the opportunity this provided for learning and relationship building within alliances.

¹ For a description of the 6 OH steps and 9 OH principles, together with rich reflections and resources, see the core text on OH: Wilson-Grau, R, 2019, Outcome Harvesting Principles, Steps and Evaluation Applications, IAP, Charlotte, NC.

1.3 Scope and structure of this report

The focus of the report is the use of OH for monitoring, not evaluation. A survey was designed to capture both the experience of organisations in using OH in D&D alliances as well as the more extensive experience of some respondents, including GPPAC, Hivos and Oxfam Novib. A mix of required and optional questions was used in the survey, hence there is some variation in the number of responses for each question. sections two-seven of the report follow the structure of the survey:

- Monitoring choices
- Harvesting outcomes
- Reviewing outcomes
- Substantiating outcomes
- Analysis and use of outcomes
- Assessment of OH experience

In section 8 of the report, highlights from two learning events are presented. The learning events were convened by the OH task team to allow survey respondents to reflect on and further examine the survey findings.

1.4 How to use the report



Quotations from survey responses are indicated in blue boxes or, when in the main body of the text, with blue text e.g.: *Quotation from a survey response*.

‘Short quotation from a survey response’

Comments by the author are indicated as follows:

The author’s reactions to findings and themes that could be further explored.

Symbols are used to highlight findings and author comments as follows:

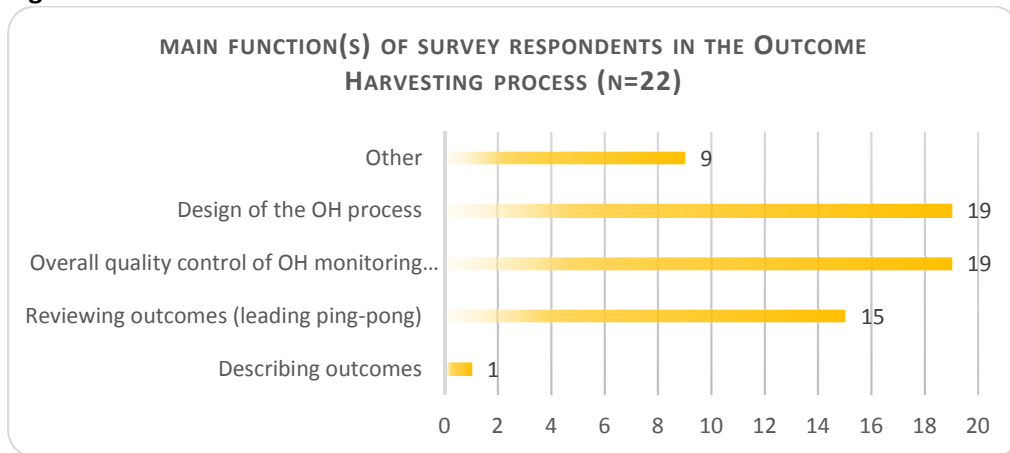
	Good practices , both those evident in the findings and those suggested by the author
	Points of interest that could be explored further e.g. in a learning event

1.5 Survey respondents

Sections 2 - 7 of this report are based on responses from all 22 organisations / alliances that were invited to participate in the survey; see Annex A. Survey respondents brought various perspectives based on their roles in the OH process. The majority of those responding (19 of the 22) were responsible for the design of the OH process and overall quality control. Survey respondents also shared experiences of reviewing outcomes and their other OH functions such as training, analysis and overall coordination and, in one case, the integration of OH into funding proposals. In the case of 12 alliances, the survey responses represented a greater diversity of views because the person submitting had been able to collate views of at least one other person in their alliance. In six cases, the consultations were particularly extensive, including in some cases those who write outcome

statements, facilitators of harvests within alliance members, facilitators of ping-pong (review) processes in country, and those who use outcomes in annual reporting.

Figure 1



In summary, the findings are based on survey responses that represent views from across the OH process. Most responses represent the collated views of at least two people in an alliance. However, the responses are not a systematic representation of the views of all actors in the OH process in all alliances.

2 Monitoring choices

Under the D&D programme, the MFA recommended but did not require the use of OH². In this section, we look at the decisions taken by the organisations and alliances that chose to use OH, or adaptations of OH, for monitoring.

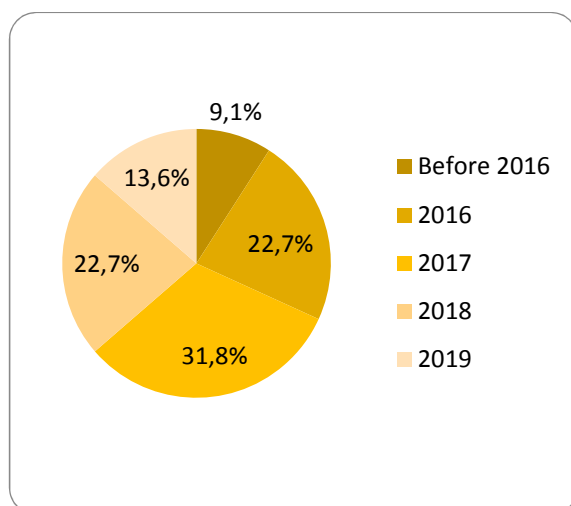
2.1 Experience of using OH for monitoring

The great majority – 91% - of organisations started using OH for monitoring after the start of their D&D grant in 2016. Two organisations have longer experience with using OH for monitoring, with one, GPPAC, having pioneered the development of OH. At least two others had experience using OH in evaluation before their D&D grants.

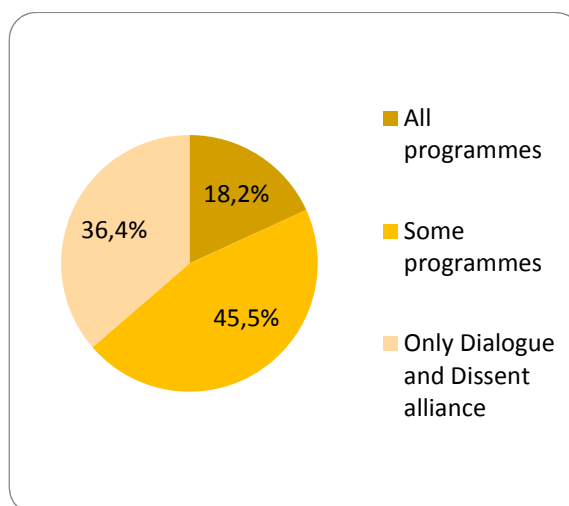
The main reason why OH was not introduced right at the start of the D&D grant in some alliances is recognition that outcomes take time to materialise so some felt it wasn't necessary to put OH into practice in the first one-two years. Alliances were also mindful of the need to first better understand the approach, some lacked a strong M&E practice / an M&E approach, others were not motivated to try OH before the mid-term evaluation.

Figure 2

When use of OH started (n = 22)



Use of OH for monitoring (n = 22)



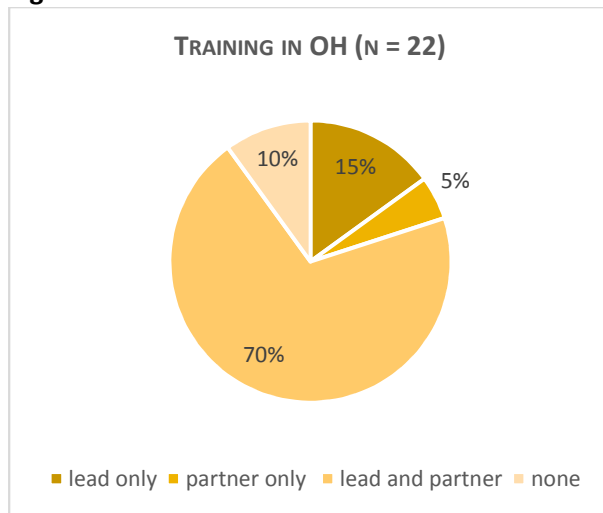
Before the D&D alliances began in 2016, only about 9% of lead organisations were using OH for monitoring. Today (mid 2020), use of OH for monitoring extends beyond the D&D programmes of about two-thirds of lead organisations. This significant expansion in the use of OH implies that the value of OH for monitoring has been recognised beyond the D&D teams.

² Version 2.0 of the Dialogue and Dissent Theory of Change (page 12) stated: *Therefore, process oriented methodologies such as outcome harvesting can be used for reporting and explaining outcomes in the area of capacity development and advocacy trajectories.*

2.2 Training

Using OH well requires, among other skills, a basic conceptual understanding of the OH/OM definition of an outcome, a clear understanding of how to formulate an outcome statement, the ability to write clearly and concisely, clear evaluative thinking and for those leading the process, excellent facilitation skills. Not everyone can develop competence in each of these skills, nor do they need to so long as overall the OH process involves people who together have these skills. Still, some level of training as well as coaching is typically valuable for those using OH for the first time and should be adapted to their role(s) in the process. Unsurprisingly, respondents indicated that training was given in 90% of alliances, most often (in 70% of alliances) to both lead and partner organisations. Of the three alliances that reported having no training, one commented that there had been training but it was quite minimal while another indicated that staff describing outcomes are trained in how to describe them.

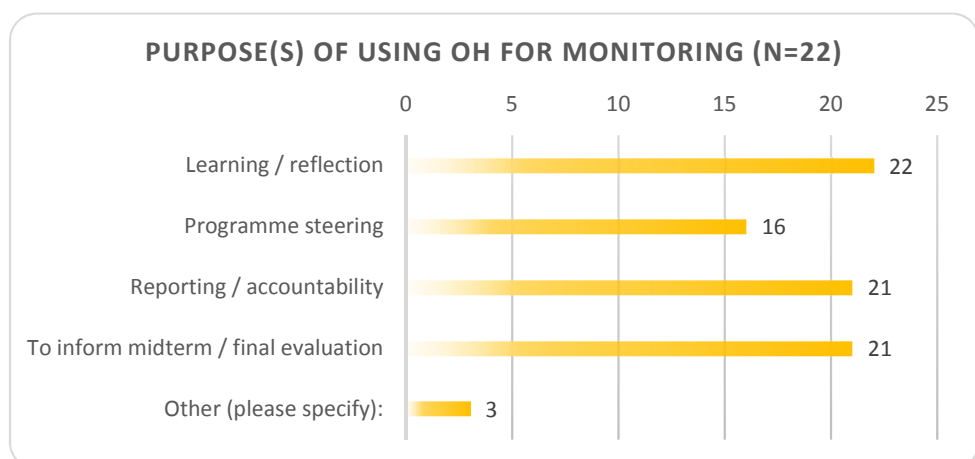
Figure 3



2.3 Purpose(s) of using OH for monitoring

In all alliances, the (intended) purpose of using OH as a monitoring approach was to support learning and reflection and in all but one alliance also to inform an evaluation and support reporting.

Figure 4



2.4 The main reasons for choosing OH

Looking at the rationale for choosing OH over other approaches, the responses largely confirm widely recognised strengths of OH. Most frequently mentioned was the **high suitability of OH for monitoring complex programming such as lobbying and advocacy as well as networks** seeking to achieve systemic changes. (The absence of good alternatives for monitoring such complex types of programming was also remarked upon.) In line with one of the commonly cited purposes of using OH, it was chosen by many to **inform adaptation of programming** through the use of harvested outcomes to **reflect on theories and processes of change**, including the effectiveness or otherwise of strategies used to achieve outcomes. Highly valued by three respondents was the **participatory nature of OH**. Only one respondent said it was being used because the MFA had recommended it, suggesting lead organisations and wider alliances felt free to decide on their PMEL approach for D&D.

[OH] 'increases ownership and common understanding of the project's objectives and results'.
SNV Netherlands / Voice for Change

2.5 Why other approaches were also needed

It is widely understood that OH cannot meet all PMEL needs. In decreasing frequency, the reasons respondents gave for using other approaches were:

- Internal programmatic reasons: Other approaches were considered more suitable for monitoring results of interventions such as capacity building, awareness raising, 'field projects' and direct poverty alleviation.
- Conflicting requirements: Incompatibility with log frame / outputs and indicators of donors other than MFA; challenges in linking with the overall result framework of the organisation.
- Capacity and resources: OH is not always feasible because it requires capacity strengthening resources to get usable quality data.

2.6 Use of the six OH steps

All steps except step 4 (Substantiation) were used by the great majority (19-21) of the 22 alliances. There were two main reasons given by eight respondents for not using substantiation: a) in some cases, it was decided to use step 4 only in evaluations; b) in several alliances there was also concern about how to do substantiation and / or the resources needed³. It is notable that there has been interest in using substantiation not only in evaluations but also when using outcomes for other reasons. See Section 5 for more on alliance experiences with substantiation.



The use of substantiation by 14 of the 22 respondents and interest in using substantiation other than for accountability / reporting is surprising. When using harvested outcomes for learning and reflection, the enhanced credibility that substantiation can bring is often considered to be less important than when using the harvested outcomes for accountability or public communication.

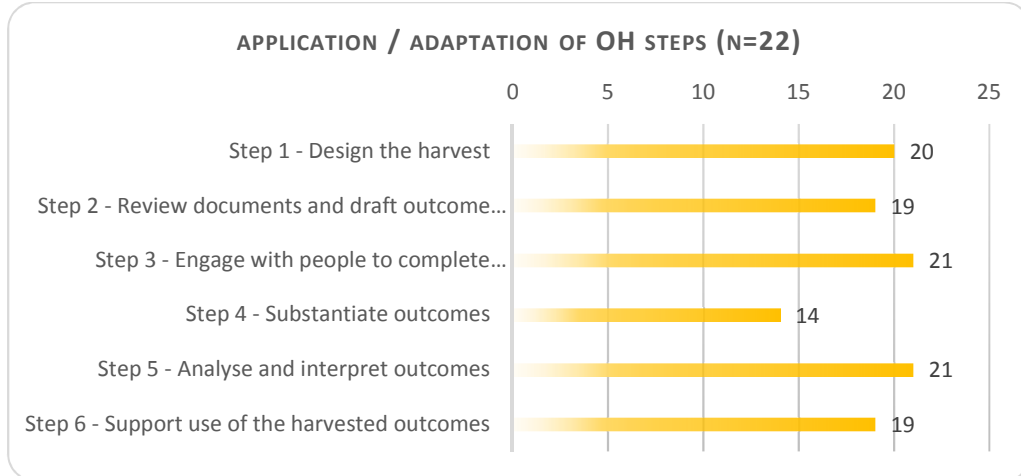
Some respondents reported using the steps in a different order or combination, such as combining the formulation and review of outcome statements with analysis in a single event or process, or returning to step 4 during an evaluation, or combining steps 5 & 6.

³ Challenges noted include what percentage of outcomes should be substantiated, how 'external' should substantiators be and how often should substantiation be done.



The use of OH steps in various combinations shows good OH practice: the steps are intended to be iterative and adaptable.

Figure 5



The standard definition of step 2 – Review documents and draft outcome statements – does not correspond closely to the use of OH for monitoring when it can be expected that the first description of outcomes is done by people knowledgeable about the outcomes without reference to already published documentation on the project in question. At first sight, therefore, it is somewhat surprising that 19 of the 22 respondents reported using step 2. Although not clear from the survey question and responses, the most likely explanation is that respondents interpreted this option as ‘draft outcome statements’ rather than ‘review documents’.



Have alliances used document review as the first step in identifying outcomes or have harvests always started with people describing outcomes directly? If both approaches have been used, what are the pros and cons?

In summary, only 9% of alliances were using OH before the D&D programme but now about two-thirds are using it for monitoring not only the D&D programme but other work as well. Training has nearly always been used and usually involved both partners and lead organisation staff. OH was chosen to support learning and reflection as often as to support accountability and evaluation. The main reasons for choosing OH were its suitability for advocacy and influencing programmes and the usefulness of outcome statements for reflecting on and adapting theories of change and strategies. As expected, other approaches are in use by all alliances to meet PMEL needs that OH cannot address well. Surprisingly, step 2 – document review – has been widely used in monitoring; it would be useful to clarify how and why this is so. Concern was expressed in how to use step 4 – Substantiation.

3 Harvesting outcomes

In OH, as with participatory M&E in general, the choice of tools and techniques used to harvest outcomes, as well as the frequency of their use, depends on the purpose(s) of the harvest, the organisational context and the individuals involved. Below, we explore some of the main features of how outcomes were actually harvested.

3.1 Frequency of harvesting

Most commonly, the frequency of harvesting has been between quarterly and annually with just three respondents harvesting outcomes monthly or in real time (ongoing). Notably, seven alliances harvested outcomes at two-three frequencies; some mentioned that the frequency depended on the country / partner. While two respondents commented that they were able to harvest successfully quarterly or biannually, several respondents have an aspiration to harvest at such frequencies to inform reflection but in reality, harvesting has been focused largely on the (obligatory) annual reporting. This is not unusual as reporting is an additional incentive with a fixed deadline and annual harvesting can, as in some alliances, sometimes be done face-face as part of annual reflection / planning meetings.

'Some teams have been able to do ongoing harvesting. Others do it twice a year when a reporting deadline is due.'
 AMREF / Health Systems Advocacy for Africa

Table 1: Frequency of harvesting (n=22)

Alliance / organisation	Annually	Biannually	Quarterly	Monthly	Ongoing
1. Shared Resources, Joint Solutions	No	Yes	No	No	No
2. Conducive Environment for Effective Policy	Yes	No	No	No	No
3. PITCH (Beat the Aids Epidemic)	Yes	No	No	No	No
4. Towards a Worldwide Influencing Network	Yes	Yes	Yes	No	No
5. Health Systems Advocacy for Africa	No	Yes	No	No	Yes
6. PRIDE	Yes	No	No	No	No
7. No News is Bad News	Yes	No	Yes	No	No
8. RNW Meida	No	No	Yes	No	No
9. Civil Engagement	No	No	No	No	Yes
10. Green Livelihoods Alliance	Yes	Yes	No	No	No
11. Advocacy for Change	Yes	No	No	No	No
12. Citizen Agency Consortium	Yes	No	No	No	No
13. Prevent up Front	No	No	Yes	No	No
14. Right Here Right Now	Yes	Yes	Yes	No	No
15. Partners For Resilience	No	Yes	No	No	No
16. Watershed	Yes	Yes	No	No	No
17. Capacitating Change	No	No	Yes	No	No
18. Building Capacity for Sector Change	Yes	No	No	No	No
19. Freedom From Fear	Yes	No	No	No	No
20. Voice for Change	No	Yes	Yes	Yes	No
21. Girls Advocacy Alliance	No	Yes	No	No	No
22. Every Voice Counts	No	No	Yes	No	No
Total	12	9	8	1	2

3.2 How outcomes were harvested

In each application of OH, those facilitating the harvest need to find the most effective and feasible means of getting those who know the outcomes to first identify and then to describe them in writing as outcome statements⁴. The responses indicate a mix of face-face and virtual approaches were used. Most common were face-face group sessions and individuals describing outcomes remotely.

Table 2: Frequency of face-face and virtual approaches to harvesting (n=22)⁵

Alliance #	Face-face		Virtual	
	Individually	Group sessions	Individually	Group sessions
1	sometimes	sometimes	often	often
2	not used	sometimes	often	sometimes
3	not used	often	not used	sometimes
4	not used	sometimes	often	not used
5	not used	sometimes	often	sometimes
6	sometimes	sometimes	often	not used
7	sometimes	often	sometimes	not used
8	not used	not used	often	sometimes
9	not used	not used	always	not used
10	sometimes	sometimes	sometimes	sometimes
11	not used	always	not used	not used
12	not used	often	often	often
13	sometimes	sometimes	always	sometimes
14	often	often	often	often
15	sometimes	often	sometimes	often
16	sometimes	often	sometimes	often
17	sometimes	often	sometimes	not used
18	sometimes	often	sometimes	sometimes
19	sometimes	sometimes	often	not used
20	often	often	often	sometimes
21	sometimes	always	sometimes	sometimes
22	often	often	not used	not used
Total - at least sometimes	14	20	19	14

⁴ The core text on OH (Wilson-Grau, R, 2019, Outcome Harvesting Principles, Steps and Evaluation Applications, IAP, Charlotte, NC) refers to those who describe outcomes as 'sources'. This term has proven unpopular with several alliances who feel that it does not appropriately reflect the participatory nature of OH. Consequently, the task force designing the survey chose not to use the term 'sources' and this report likewise avoids the term in favour of the arguably more neutral, if somewhat cumbersome, 'those who describe outcomes' or similar.

⁵ See Table 1 for the names of alliances / organisations

Very few – just two – respondents always used just one particular approach for engaging with those describing outcomes; instead, most alliances have used a variety of approaches indicating that they needed to be flexible and potentially opportunistic in how and when outcomes were harvested. In one case, the respondent did not know how partners harvested. For others, most outcomes were harvested using Word or Excel reporting templates. Several commented that the ‘pressure cooker’ approach of using writeshops has been effective. In some cases, hybrid approaches were used, for instance:



- Local partners record draft outcomes on a form prior to quarterly reflection meetings where they then select, revise and validate outcomes.
- Individuals in CSOs note outcomes in their logbooks; further outcomes are identified in consultation with project teams; CSOs also jointly identify outcomes during learning / capacity building events.

With the new COVID-19 realities, those who have relied on face-face harvesting realise they will need to adapt and use virtual approaches. While not ideal in terms of engagement, this should be feasible: virtual approaches are already as commonly used by alliances as face-face approaches.

Turning now to the use of tools or software in harvesting (initial data collection, not reviewing), all but one respondent indicated the use of some kind of tool to facilitate the process. In about 15 alliances, tools were limited to forms created in Word or Excel or project management software. These were either integrated with other monitoring templates or used separately. Some mentioned providing guidance. Only two mentioned specific tools, one using the OH App from Kwantu, another a custom tool made using Kobo Toolbox. Two practices used by alliances are particularly noteworthy:



- Some have outcomes entered directly into Excel or a custom database where the person describing the outcome also does basic analysis to classify the outcomes.
- Others favour low-tech but motivating approaches: warm up exercises to shift thinking from activities to achievements, then use of flip charts to draft outcome statements followed by data entry. Another example: timeline reconstruction is used with beneficiaries and partners to identify outcomes.

3.3 Lessons learned

Survey participants were asked to identify the parts of the harvesting process they would maintain going forwards and those they would like to do differently. Of the 22 responses, 15 identified aspects of the harvesting that they would maintain⁶. This is an encouraging indication that respondents have valued experiences to share and build on. They also feel they have much to learn: all but one respondent identified changes they would make and overall the number of ideas or aspirations for improving harvesting outnumbered the aspects to be maintained. Three main themes are evident in responses:

- Process
- Capacity
- Tools

‘The narrative character of OH makes it very useful information for PMEL & project staff alike.’
Milieudefensie / Green Livelihoods Alliance

3.3.1 Practices to maintain



The following practices and processes most valued by respondents all resonate with the author’s experience.

⁶ The survey did not require respondents to answer both what to maintain and what to change; accordingly not all respondents identified both.

Most respondents expressed satisfaction with many aspects of the process they have used.

Around half of respondents mentioned that the frequency and mode of harvesting has worked particularly well. Notable examples of processes valued by respondents include:

- **Combining short cycle of reporting with quarterly physical meetings** for joint learning and strategizing.
- **Starting workshops with a brainstorming phase**, then decision making on who elaborates which statement prior to embarking on the writing of the outcome statements. This was said to help in celebrating accomplishment.
- **Integration of harvesting and reflection** on, for instance, a theory of change in a single event or process increases partner ownership and common understanding e.g. of the role of the project in the bigger picture.
- **Being opportunistic**: seizing opportunities for harvesting with partners as they arise.

'The write-shop model has become popular among the harvester coordinators for formulating outcomes.'

Hivos / Citizen Agency Consortium



Some highlighted how they had been successful in maximising engagement in the harvesting process through, for instance, using small groups to overcome a reluctance of some to be vocal in larger groups. Such comments reflect the importance of strong facilitation of the OH process.

Two respondents who had used OH with another approach (MSC and historical reconstruction) were positive about continuing these processes.

Next in terms of frequency of comments were those relating to good practices in **capacity strengthening**. The value of learning by doing, one of the nine OH principles, was recognised by some. Ongoing training as personnel change or to support the delegation or sharing of responsibilities with local or partner staff has been valuable. One respondent reported success in embedding funds for OH capacity strengthening in funding proposals.

3.3.2 Practices to change

Above all, ideas for changing the harvesting process concerned the question of **how to improve outcome quality**. Some plan to give more time to initial descriptions or later improvement. Decoupling the harvest process from reporting and payments to partners was suggested as a way to give more time to the harvest process. Another cautioned that final outcome statement quality could be stronger if the significance and contribution were elaborated when first describing outcomes rather than to try to do so at a later stage as previously. The potential for more peer review / internal quality checks by lead organisation staff or partners was noted. Set against these aspirations for greater quality, two respondents noted the challenge of achieving quality data while wanting light or at least feasible processes.



Peer review of outcome statements potentially has three advantages: suggestions for improvement to outcomes, increased capacity of those doing the peer review, and knowledge sharing through familiarity with outcomes beyond your immediate area of work.

Two respondents indicated an interest in **using OH from the outset**, either by starting OH immediately when a project begins or involving partners in the harvest design phase.



Using OH from the outset of a project brings the benefits of a participatory reflection process into a project from the start; involving partners in the harvest design is good practice for ensuring the considerable efforts involved in OH meet the needs of partners as well as lead organisations.

Some are keen to explore how variations to the **frequency of harvesting** can improve harvesting. Three respondents want to try harvesting in real time or at least every two weeks. Another sees value in biannual identification of outcomes by smaller teams to reduce the burden of annual harvesting.



In principle, with more frequent harvesting it should be easier / quicker to recall information and less chance of forgetting important outcomes. Balanced against this is the challenge of engaging people to describe outcomes more frequently. Where it is easier to engage some people more than others, the harvested outcomes may not then reflect the full range of perspectives and knowledge in an organisation or alliance. A possible middle ground is to 'harvest light' more frequently / continuously, from which possible outcomes are identified. This can be followed by a selection and reflection process – perhaps in a writeshop - involving in developing full outcome statements.

Three respondents reflected on **the best ways to give voice** to all those needed to make a particular harvest informative and useful. More joint harvesting of outcomes and interactive writeshops from the beginning of the programme were two ideas. Set against this are the challenges of **managing power relations**, as when government officials are involved in a workshop, or with NGOs wanting to claim leadership rather than contribution.



A practical suggestion with considerable merit from one respondent concerns how to manage power relations: involve those whose views could dominate in only a part of the harvest process when they can contribute new outcomes and / or validate outcomes formulated by those who are less powerful.

Lastly, four respondents noted their wish to **integrate harvesting into reporting and / or planning processes**, presumably to increase engagement, efficiency and use of findings.

Almost as numerous as ideas to improve the harvesting process **were aspirations for strengthening capacities**. The capacity needs fall into the following clusters:

- More training for some partners in order to improve their identification and description of outcomes
- Training when new staff arrive / people move on

- Refresher training
- Facilitation training for staff, partners and consultants because good facilitation is key to good OH
- A larger PMEL team to support the process across alliances

Several respondents are looking for a single tool to facilitate harvesting, one that provides a single database, accessible to partners, that allows easy posting of outcomes, facilitates ping pong, supports at least basic classification of outcome data and allows supporting evidence to be linked or attached to submitted outcomes. One has identified Podio⁷ as a candidate to be evaluated.



Podio is worth exploring as a tool to facilitate harvesting as it meets all the above criteria. I have found it works well when multiple people are involved in harvesting. Crucially, its commenting and editing feature is useful for ping pong.

For two alliances, there is a strong interest into how to assess progress with capacity strengthening. One respondent has found OH to be moderately useful but wants to combine it with other approaches, another is less sure if OH could be useful or if other approaches are preferable.



What lies behind the varied experience with using OH for assessing capacity strengthening? From my experience, I have found that OH can be a powerful way to evaluate capacity development where changes in behaviour of an actor can reasonably and plausibly be linked to activities of an intervention, as when an intervention introduces a new approach or tool.

⁷ <https://podio.com/>

4 Reviewing outcomes

4.1 How outcomes have been reviewed

All 22 respondents indicated that outcome statements are reviewed before being accepted and used, however six indicated that the process was not ‘formally defined’ or was somewhat variable in how it was implemented. The descriptions of the review processes indicate considerable variability, a potentially good sign that processes have been adapted to organisational contexts. Of some concern, however, is that responses from some alliances suggest that there has not always been strong engagement between those coordinating the harvest in country or in the Netherlands and those describing outcomes. Typically, the ping pong / review process is critical in OH as first submissions of outcome statements are rarely sufficiently clear or SMART⁸ to be credible enough for any uses. Of course, when using OH for monitoring the competence of those describing outcomes can improve over time, resulting in a tapering need for significant clarifications or changes. But such progress is usually variable and many people never become strong at formulating outcome statements, hence the critical ongoing need for a review process.

Considerable variations are evident in the review processes used across alliances, as illustrated by the following summarised examples:

- **Two-step process – in country and Netherlands:** In-country harvester reviews first, followed by peer review with team, then submit to PMEL of the Netherlands NGO, then ping pong with informants as necessary. PMEL of the Netherlands NGO signs off.
- **Light touch first:** Simplified reporting of outcomes (not significance and contribution) reviewed by Netherlands NGO PMEL, discussed with partners each quarter. Selection of outcomes developed into statement. Then reviewed with PMEL officers. Then uploaded to Netherlands NGO’s system.
- **Peer review in writeshop** after earlier reviews by OH coordinator (in country) and headquarters. Final sign off (in theory) by headquarters during analysis for reporting.
- **First peer review in groups**, then remote team review, final review by PMEL.
- **Unclear process locally:** for example, ‘partner and country office develop descriptions’ or ‘country PMEL compiles and submits quarterly’. Netherlands NGO staff review and lead ping pong. Project Manager and / or PMEL sign off.
- **No fixed process:** Peer review during workshop. Facilitator asks questions. If possible, revisions. No ping pong after. Sometimes clarification to focal points if unclear but not edited.

The process described by one organisation shows how a review process needs to be nuanced according to where in a multi-level organisation or alliance an outcome description originates – country level, regionally or globally. At the country level, PMEL team or programme personnel give feedback or themselves suggest modifications of draft outcome statements to informants. Members of a wider consortium are then engaged to review statements, sometimes in a virtual meeting or via email. For outcomes at the regional and international level, relevant consortium members contribute to the outcome (or one consortium member if it was the only one involved in the outcome). Then a feedback process between the lead agency and the writer(s) follows. The lead agency has the final sign off on outcomes.

In general, **no particular software has been used for reviewing outcomes**, except in three alliances that used Kwantu’s OH app⁹, a custom database, or the ProjectConnect platform from Matthat

⁸ <https://outcomeharvesting.net/outcome-harvesting-smart-me-outcomes/>

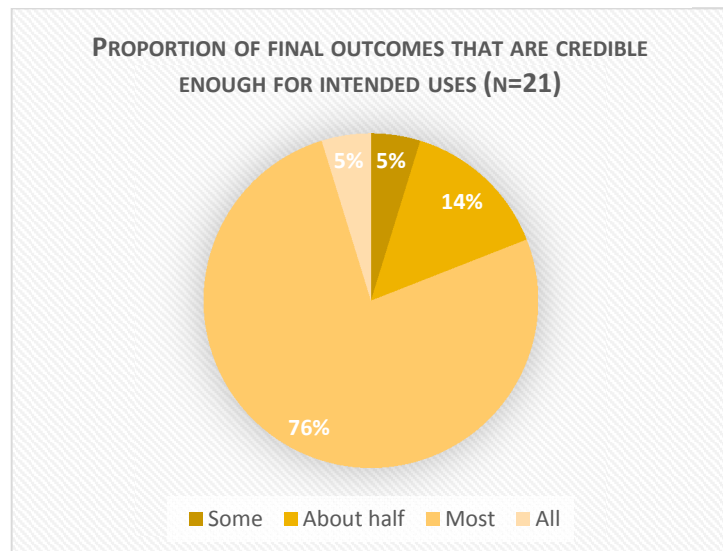
⁹ <http://www.kwantu.net/outcome-harvesting-app>

Software¹⁰. As mentioned above, Podio is another option and one that has commenting / track change feature that could be used to facilitate outcome review while maintaining the outcome statements in a database accessible to all users.

4.2 Quality and completeness of final outcome statements

After the review process, outcome statements need to be fit for purpose i.e. their quality needs to be sufficient such that they are credible enough for intended uses. Looking at how satisfied respondents were with the quality of outcomes, we see that 81% found most or all the outcomes were credible enough for intended uses. Considering that all but one respondent indicated their intended uses of OH included accountability / reporting, then either alliances have been highly successful at mainly harvesting fully SMART, i.e. high quality, outcome statements, or some outcome statements were judged to be 'good enough' but were not fully SMART. There are indications in the survey data that the latter approach may have been followed by some: more than one respondent indicated the effort to develop SMART statements has been concentrated on a selection of outcomes. See sections 5 and 8.1 for more on outcome quality.

Figure 6



Have alliances been equally satisfied with the quality of outcomes for internal vs external uses? Were lower quality standards accepted when using outcome statements internally?

'two-thirds of outcome statements were SMART enough, [but] all were useful for sensemaking.'
IRC / Watershed-Empowering Citizens

¹⁰ <https://projectconnect.info/>

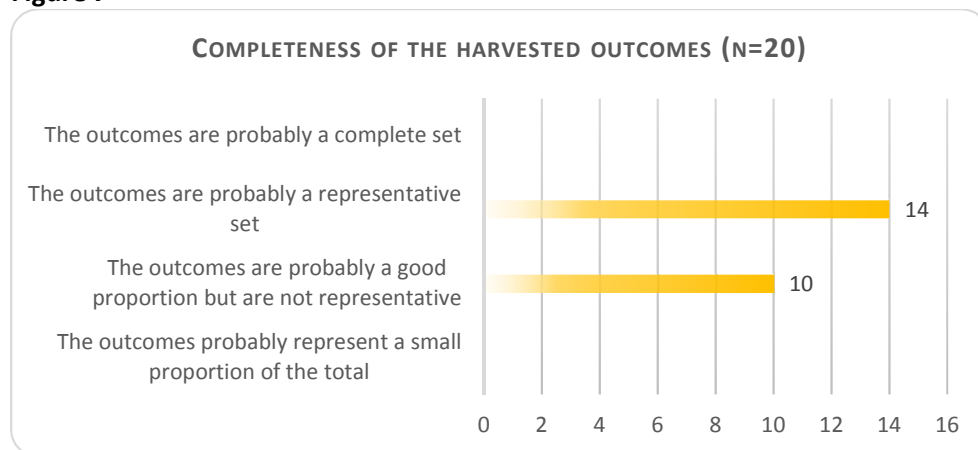
The survey didn't tell us if outcome statement quality differed for data used internally versus externally but did indicate that the overall quality of final outcome statements was far from uniform. The limited description of contribution was recognised by some as a particular weakness of outcome statements. Variability was noted between projects, those describing, countries, partners, levels in organisation, years and from different translators where the original description was in another language.

It is not unusual for contribution descriptions to be vague or lacking detail e.g. dates initially. It is also not so uncommon for initial contribution descriptions to include activities that occurred after an outcome. Patient but persistent ping pong is needed to ensure contribution descriptions are clear and plausible.

'Quality and level of detail varies though, this is obviously a challenge with 14 programmes and over 50 organisations involved, staff changes etc. - especially if you value local ownership.'
Plan Netherlands / Girls Advocacy Alliance

Respondents were asked to indicate the extent to which they consider the harvested outcomes represented all of the outcomes achieved by the monitored projects. Of the 20 out of 22 alliances responding, all considered the harvested outcomes to probably be either a good proportion or a representative set of all the outcomes that actually materialised. In general, respondents said that more outcomes could be harvested and that it was not a question of involving different people, rather more time would be needed with those already engaged. Some respondents noted that the extent to which a set of outcomes was comprehensive depended on the types of the outcomes; this resulted in more than one response to the survey question. Interestingly, the types of outcomes which were more comprehensively harvested differed between alliances: for one, advocacy-related outcomes were considered to be representative but capacity building outcomes not so; in another alliance the opposite has been the case. Another variation is that for one alliance the more important advocacy outcomes have been described but not those that are more minor.

Figure 7



Another factor influencing the completeness of harvests was the **frequency of harvests**. Less frequent harvests such as once a year were said by some to probably be a limiting factor as participants in the harvest are likely to have forgotten some outcomes.



An approach used with reportedly great success by one alliance is very frequent harvesting tied to short-cycle reporting and reflection. Even if an outcome is missed in one report, if it was important it would come up in the next quarter, hence over a year harvests are considered to contain all significant outcomes.



Three respondents indicate a further reason harvested outcomes were not fully representative: there were **few if any negative or unintended outcomes harvested**. In contrast, some alliances reported notable success in harvesting such outcomes – see section 7.1.



Harvesting unexpected or negative outcomes is commonly a challenge. Steps that can be taken to try to encourage the description of such outcomes include making safe learning spaces for harvesting and / or dedicating part of the harvest process (workshop, interview, form) to such outcomes.

Overall, the responses indicate that the monitoring processes have been as effective as they can reasonably be expected to be given the D&D and other projects featured in the survey responses typically involved many actors / are large scale and that the review of outcomes in OH is time intensive. That said, what is not so clear in the responses is the actual objectives of monitoring and the extent to which these were met.



Was monitoring intended to be comprehensive or was it intended to capture the most important outcomes and, if so, how successful was that?

When a harvest cannot be comprehensive, it should still follow a systematic approach in order to be consistent over time and between locations. Did alliances use criteria – e.g. aim to harvest the most significant outcomes and define ‘significance’ – to prioritise their harvesting efforts and, if so, how well did this work?

4.3 Lessons learned

4.3.1 Practices to maintain

Several respondents were satisfied with the review process as it stands. Notable themes in the comments are:

- Keep focussed on what is ‘good enough’ for purpose and clarity for an external reader. There is no need to perfect all outcome statements; when you come to use them in e.g. reporting, then you can add missing information.
- Ping pong can be demotivating or draining. Maintain interest by adopting more appealing ways for ping-pong (meetings/ skype contact) instead of only email.
- Combination with MSC during reviewing means going into depth on the outcomes; by formulating the story behind it, we get much more information than through a ping-pong exercise.

4.3.2 Practice to change

Reviewing / ping pong of outcomes is time-intensive and often accompanied by time pressure of reporting. The challenges this has caused respondents is evident in the ideas for making the review processes more efficient. Starting with the most frequent, the main themes are:

- Use a more structured process / protocol for giving feedback, improving statements and collecting evidence, additional outcomes, sign off
- Invest in coaching and support so a wider pool of staff can lead ping pong, take ownership of the quality control process and learn to formulate outcome statements better.
- Consider Podio or other software to facilitate the review process.
- Reorganise the harvest / review process to allow more time for reviewing, e.g. by having regular or real time submission of outcomes.
- Use peer review of outcome statements within an alliance.

5 Substantiation

Step 4 in OH, Substantiation – asking people independent of an intervention to confirm the accuracy and / or comment on an outcome statement – is one way of confirming that outcomes are accurate and hence credible enough for their intended uses. We have seen above that eight of the 22 respondents have chosen not to use substantiation so far¹¹. Before examining how it has been used, let us put its use in context by understanding what other measures have been used to enhance the credibility of harvested outcomes across the 22 alliances. The table that follows indicates the frequency with which each alliance has used six measures, i.e. actions intended to produce an effect, in this case an increase in data credibility¹²:

- **Agreement on who describes outcomes.** Primary users agree with the evaluator / harvest coordinator on reliable primary and secondary sources.
- **Application of the OH SMART criteria.** Outcome statements that are SMARTly formulated are, among other qualities, observable and hence verifiable.
- **Description of outcomes by teams** so that each outcome statement represents a collective rather than an individual assessment of the facts of the outcome and the intervention's contribution as well as a collective view on the outcome's significance.
- **Internal verification of outcome statements**, i.e. confirmation of the accuracy of outcome statements by people internal to the intervention whose outcomes are being harvested. Such people need to be knowledgeable of the outcomes but not to have been involved in the original formulation of the outcome statement(s) they review.
- **Attribution of outcome statements** to those who formulate them, also known as 'going on record'.
- **Substantiation of the accuracy of outcome statements** by one or more people who are independent of the intervention whose outcomes are being harvested.

Some respondents mentioned a further measure that has been practiced by some alliances: triangulation of outcome statements with external or internal documents. The extent to which such triangulation has been used is not evident in the survey data as there was no specific question on triangulation, rather it was assumed that responses on use of substantiation would capture triangulation with documents.



Triangulation of outcomes using supporting documentation as practiced by some alliances is a way to enhance credibility.



Arguably, simply requiring those describing outcomes to identify supporting documents and / or possible substantiators in the knowledge that such sources may be consulted is in itself an additional credibility check, even if only a proportion are used for triangulation.

¹¹ That eight out of 22 alliances reported not using substantiation so far implies that 14 alliances have used substantiation. However, responses on the reasons for choosing substantiation were provided by just 13 respondents.

¹² The six measures that can increase data credibility in OH are derived from the discussion on data credibility on pages 88-89 in Wilson-Grau, R, 2019, Outcome Harvesting Principles, Steps and Evaluation Applications, IAP. Internal verification is not explicitly mentioned in this reference but refers to the review of outcome statements by other people internal to the intervention whose outcomes are being harvested. Such people need to be knowledgeable of the outcomes but not have been involved in the original formulation of the outcome statement(s) they review.

Table 3: Frequency of use of data credibility measures¹³

Alliance #	Agreement on who will describe outcomes	Application of the OH SMART criteria	Description of outcomes by teams	Internal verification of outcome statements	Attribution of outcome statements	Substantiation
1	always	often	sometimes	sometimes	unsure	sometimes
2	always	always	sometimes	often	always	not used
3	always	unsure	often	always	sometimes	not used
4	often	always	sometimes	sometimes	sometimes	sometimes
5	sometimes	sometimes	sometimes	not used	not used	not used
6	often	sometimes	often	often	always	sometimes
7	unsure	often	sometimes	always	often	often
8	always	always	often	sometimes	not used	sometimes
9	unsure	unsure	sometimes	sometimes	unsure	sometimes
10	unsure	always	unsure	unsure	always	not used
11	always	often	always	always	unsure	always
12	unsure	always	sometimes	not used	always	sometimes
13	always	always	sometimes	sometimes	always	sometimes
14	often	often	often	often	often	not used
15	often	often	often	sometimes	sometimes	not used
16	always	often	always	often	sometimes	not used
17	often	often	often	sometimes	not used	not used
18	unsure	not used	often	often	not used	sometimes
19	unsure	unsure	sometimes	often	not used	not used
20	always	often	often	always	often	sometimes
21	unsure	always	often	unsure	unsure	sometimes
22	always	not used	often	sometimes	often	sometimes
Total - at least sometimes	15	17	21	18	13	13

We can see from the table that there was considerable variation in the measures taken by alliances to ensure data credibility. All alliances used two or more credibility measures but in many cases there was considerable variability in the measures used: some 21 alliances used various measures ‘often’ or ‘sometimes’ and only one alliance always used the same measures. Description of outcomes in teams was the most commonly used measure: it was used by all but one alliance. Notably, five respondents reported that SMART criteria either hadn’t been used at all or they were unsure about their use. Also notable is widespread uncertainty about whether some of the other measures had been used. One reason for this could be that the survey did not include definitions for each measure as it was assumed there was a common understanding of them. **Outcome quality and credibility measures are examined further in section 8.1.**

¹³ See Table 1 for the names of alliances / organisations



The OH SMART criteria are probably the most widely accepted way to ensure the quality of an outcome statement. I find it is surprising, therefore, that five alliances did not use these criteria or respondents were unsure of if the criteria had been used. Also notable is the widespread uncertainty about the meaning of other credibility-enhancing measures. Does this show the survey lacked sufficient explanation of these measures or that several respondents were not knowledgeable about them?

Turning now to the 13 alliances that have sometimes used substantiation, we look first at why they used this OH step. As expected, the main reason for substantiation was to use outcomes for reporting / accountability. Often comments specifically mentioned substantiation was only done to enhance the credibility of monitoring data used in mid or end-term evaluations. More surprising is that outcomes were also substantiated for use in learning and programme steering. This shows the importance for some alliances of substantiating outcomes even for internal use.

Figure 8



The substantiation of outcomes for use internally for learning and reflection and / or programme steering is somewhat surprising considering it is most commonly used when sharing outcomes externally. What were the considerations informing the decisions to use substantiation for internally-used outcomes and do all respondents intend to continue with this practice?

Notably, the three alliances that most often chose to use substantiation to deepen their understanding of outcomes also used substantiation for learning or programme steering. For 10 of the 13 alliances, substantiation did bring new insights into outcomes but this was a side benefit, not an objective of substantiation.

5.1 How outcomes were substantiated

In all uses of substantiation, a sample of outcomes rather than all outcomes was selected. There was considerable variation in the criteria used to guide the selection of outcomes for substantiation. In one case the selection was random, in another there were no criteria and for one alliance the sole criteria was not to compromise safety or confidential details of partners and communities. More commonly, a set of criteria have been applied, as in the examples that follow. A common but not universal criteria was the selection of outcomes for which other supporting evidence was not available or was weak.

Figure 9

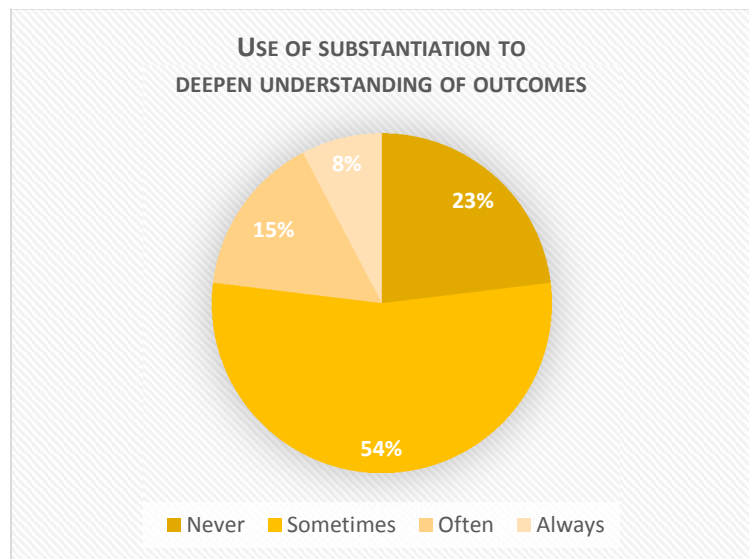


Table 4
Example criteria used to select outcomes for substantiation

<ul style="list-style-type: none"> • Most significant in each country context • Weakest evidence / least credible • Importance to objectives • Informative for understanding process of change • Contribute to learning agendas 	<ul style="list-style-type: none"> • Relevance to ToC • No other evidence • Uncertainty about outcome, contribution, or contribution of others • Feasibility of finding substantiators 	<ul style="list-style-type: none"> • Uncertain / marginal contribution • Unclear significance • Differing internal opinions on significance / contribution • Most recent of a sequence of related outcomes
--	--	--



Together, the criteria used by various alliances to select outcomes for substantiation provide a rich set of options. The most appropriate criteria to use will vary with each application of OH.

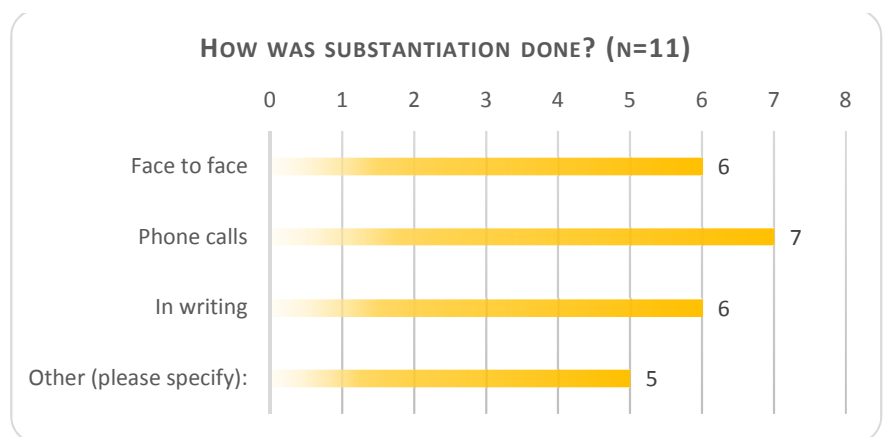
The 13 respondents who have used substantiation reported that in 16 applications of the step, **nine have been led by staff and seven by consultants**. In one alliance, substantiation was done internally out of choice. The reason for staff leading the process in other alliances is not clear from the responses. This is interesting in that some argue that externally-led substantiation is more credible because it is more independent.

Identification of substantiators is widely recognised to be sometimes challenging in OH because they need to be people knowledgeable about not only the outcome but the contribution, while also not having responsibility for delivery of the project. Encouragingly, responses did not suggest that the identification of substantiators has been an overwhelming obstacle. Some challenges were identified, however:

- Confidentiality of outcomes / need to avoid advocacy targets so as not reveal strategies and objectives. This is a common challenge in substantiating advocacy outcomes.
- When institutions rather than individuals have changed and there is no clear person who is knowledgeable of all aspects of the outcome statement.

Engaging substantiators has often involved a mix of approaches: face-face, calls and emails. Messaging was also mentioned. This range of responses confirms that the best approach depends on who you are trying to reach: some people don't have or rarely respond to emails, for instance, while others are excellent email correspondents.

Figure 10



A choice to be made when doing substantiation is how many substantiators to approach. The 10 alliances providing information have approached between one-three substantiators for each selected outcome with two-three being most common. Next, you need to consider how many positive substantiator responses are needed to confirm an outcome. Do all have to confirm the outcome, or only one or two? Only three of the 10 alliances providing information had a criteria for how many substantiators responses were needed.

A further option is to consider that substantiation of a sample of outcomes is sufficient to confirm the full set of outcomes from which the sample was derived. If choosing to do this, then two decisions are needed: first, how to sample the outcomes – both the number of outcomes and sampling criteria – and, second, the threshold i.e. the percentage of sampled outcomes that need to be confirmed by substantiation to impart confidence that the whole set are credible enough. Of the 10 relevant survey responses, seven indicated that such considerations were not used in substantiation, two responses were unclear and one reported initially using a 30% sample with a 90% threshold, then switching two years later to not setting a percentage sample size nor any threshold.

5.2 Lessons learned

The overwhelming sense from survey responses is that even those alliances that have tried substantiation feel they have a great deal to learn and almost without exception are hesitant to suggest what they would maintain or change going forwards. Two overarching challenges are evident in the following comments:

We are still trying to find the balance between substantiation for accountability purposes and a monitoring/learning/adaptation tool. Substantiation seems to deliver limited extra information but comes with significant time and budget investment. Oxfam Novib / Towards a Worldwide Influencing Network.

An unresolved dilemma is the question "when" to substantiate in monitoring. The downside of annual substantiation is the huge amount of time involved, the downside of substantiation only at the end is that your sample will include 'old' outcomes for which it may be difficult to find substantiators (even the harvesters may have moved on).

Although there is considerable uncertainty about this step, some alliances did identify practices to maintain or change.

5.2.1 Practices to maintain

Respondents were able to identify just one practice to be maintained, the low number further confirming the uncertainty about if and how to proceed with substantiation: select [the] most important outcomes to be formally substantiated by external evaluators, while smaller outcomes are validated through internal processes.



5.2.2 Practices to change

There was no dominant theme in the suggestions made. In line with the 'practice to maintain' just mentioned above, one respondent intends to start distinguishing the quality of outcomes needed for learning versus accountability:

Make clear differentiation between outcomes for reflection, learning and discussion on direction of programme; and for accountability.

Another idea concerned addressing credibility of outcomes on an individual basis, not in large numbers and not to consider that substantiating a sample of outcomes will be statistically representative.



Survey responses are insufficient to understand why substantiation of samples of outcomes to enhance confidence in all harvested outcomes has been very rare among alliances. This practice is commonly found to be useful in evaluation. Even if it doesn't provide statistical confidence, can confirming a sample of outcomes through substantiation still confer greater confidence in the accuracy of the full set of outcomes from which the sample was taken?

For one alliance, it will be important to have more than two substantiators for each outcome 'to make it more representative'. This suggests an interest as much in different perspectives on significance or enriching understanding of the outcome or contribution as in triangulating accuracy.

A final and creative idea is to consider conducting substantiation as a normal stakeholder engagement exercise together with the CSO partners at more regular intervals. This can lead to: ...in depth discussions with partners in a group setting and the review outcome statements based on the stakeholder perspective. This would shed light on how the outcome and the project's contribution is perceived by stakeholders involved in the process and give insight in how the programme could improve its performance.



While there is merit to the idea of involving wider stakeholders in regular harvests to gain wider perspectives on outcomes, a potential pitfall would be that stakeholders may be reluctant to be really open with project implementers about the outcomes and contributions in a face-face setting. In some circumstances, an independently conducted substantiation, with the option for anonymity, can provide the safe space needed for more open stakeholder responses.

6 Analysis and use of harvested outcomes

6.1 Who participated in analysis of outcomes

As OH is a participatory approach, some involvement of harvest users in analysis is to be expected and is desirable. In practice, participation in analysis can vary. In an external evaluation, analysis may or may not involve users to a significant extent depending on the motivation of the users and the extent to which an independent interpretation is required. In contrast, monitoring is an internal organisational or programmatic process so will inevitably involve users whether the HQ PMEL team, wider PMEL team, programme managers, country offices, CSO partners or others. Indeed, it is the engagement of such harvest users in analysis and / or reflection that can bring the benefits of OH-inspired learning to those involved in the harvest as well as a wider group of actors who neither described outcomes or coordinated a harvest. Who is involved may influence the accuracy and ownership of analysis and hence the subsequent use of findings and the underlying data. Let's explore who was involved in the alliance applications of OH using survey responses from all 22 alliances.

Responses from 10 alliances clearly indicate that the process of group reflection is typically motivating and that analysis has been undertaken at two or more levels: in country, at the alliance or global level and, in some cases also at the Netherlands or regional levels. In contrast, in five alliances analysis has only been done at either the country level or at the global / headquarters level. Other responses were ambiguous about the level at which analysis has been undertaken.

Analyses that are undertaken by actors at different levels are, arguably, more likely to be seen to be accurate by more users.

Another pattern evident in the responses is the extent to which the analysis processes have involved PMEL staff¹⁴, technical staff (e.g. programme or project managers) or a combination. In the great majority of alliances, both technical and PMEL staff have been involved in the analysis process. In some alliances, this has been at different scales with technical staff of lead and partner organisations involved in country, while at headquarter level only PMEL staff have been involved. More commonly, however, alliance-level analysis has involved both technical and PMEL staff. In at least two cases, the PMEL staff undertake a first analysis and technical staff then become involved in sensemaking.

In some alliances, partners / CSOs have also been involved, for instance through in-country sensemaking or their membership of an organisational oversight structure. In three alliances, analysis has involved only PMEL staff and in a fourth case PMEL and communications staff.



Overall, it is clear that most alliances find there is value in involving both PMEL and technical staff and to have at least two levels to the analysis process: in-country and global.

¹⁴ 'PMEL' is used here as shorthand for all variations of MEL functions used in the alliances.

6.2 Use of software in analysis

The great majority of alliances have relied on spreadsheets (typically Excel) for analysis. One reflection on their usefulness is worth highlighting as it implicitly distinguishes between the function of a spreadsheet and the sensemaking / interpretation that this facilitates:

...it is a good starting point to see in which domains most outcomes have been achieved, or to quickly find all outcomes relating to e.g. gender equality. But most analysis is a result of a more qualitative process. Milieudedefensie / Friends of the Earth Netherlands / Green Livelihoods Alliance.

Other applications mentioned were Power BI (said to be useful at an organisational level but not yet at team level), Maxqda (about which staff had mixed feelings) and Plectic (an online, collaborative diagramming tool used by an alliance to map actual pathways).

One alliance indicated an interest in testing qualitative analysis software such as Dedoose or Nvivo to 'identify common topics in stories'.



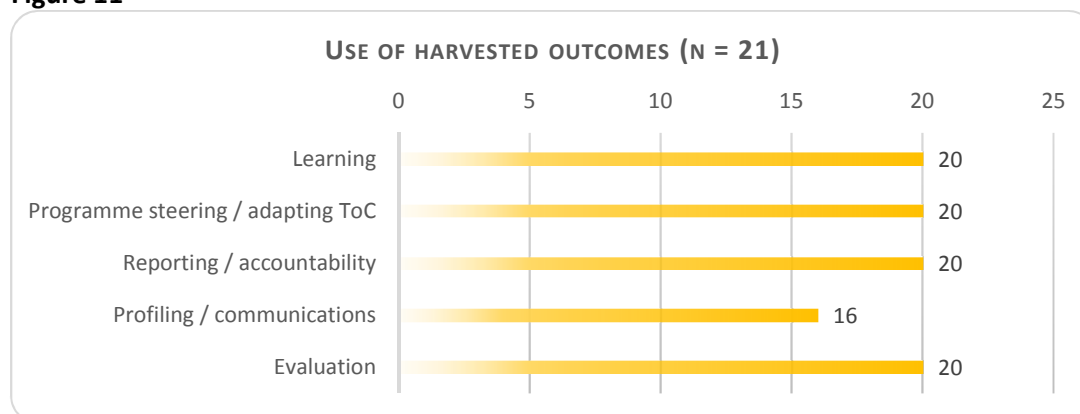
Outcome statements by definition organise data into outcome, significance and contribution, so qualitative analysis software is not helpful for this. However, there is potential utility in using qualitative analysis software with large datasets, especially those with longer outcome statements / longer elements of statements, though consideration should be given to whether and how to broaden involvement in analysis beyond the limited group of analysts who would directly code the data.

6.3 Use of harvested outcomes

Actual use of harvested outcomes is largely in line with the intended purposes of using OH defined above – see section 2.3 - but responses do indicate that more alliances have been able to use outcomes for programme steering than they anticipated (20 alliances against 18 who identified it as a purpose) and also that 16 alliances have found harvested outcomes to be useful in profiling / communications.

Reporting / accountability has been a, or the, major use of outcomes for many. In one alliance, outcomes have been the basis for quantitative reporting while also informing the narrative. Others have used outcomes to share stories of change, illustrate achievements and describe the contributions of the project to the achievements. While one alliance includes outcomes in their IATI (International Aid Transparency Initiative) report, another specifically does not do so. Reporting using outcomes went beyond donor reporting: one alliance explicitly mentioned using outcomes when reporting to stakeholders.

Figure 11



Learning and programme steering are, as one respondent remarked, closely linked. Some expressed an aspiration to strengthen their use of outcomes for these objectives by using outcomes more, more widely in the organisation / alliance, or more consistently. A challenge of going further in using outcomes for programme steering is exemplified by one response that noted that generalisations are difficult in complicated, multi-dimensional programmes. Still, many indicated successful use of outcomes in this area, highlighting for example the sharing of successful strategies, joint strategy setting with stakeholders, reflecting on and adapting ToCs and team discussion.

There has been much use of outcomes in **communications**. Uses have varied widely and commonly included the use of outcomes as the basis of evidence-based stories on website, blogs and progress reports. Initiating partnerships and evidence of achievements to cite in funding proposals were also mentioned as uses of outcomes. One noted the importance of substantiating outcomes when used for communications and the fact that some outcomes may be too sensitive to share.



Several alliances have made databases of outcomes available for external **mid-term or end-term evaluations**. No comments were made about the extent to which evaluators have been able to use the data and / or the internal analyses of the data.

6.4 Lessons learned

As with other aspects of OH use, the 22 respondents had several ideas or aspirations for further developing their analysis and use of harvested outcomes. **The responses again give an impression of a strong interest in further experimentation and development of practices.** First, let's look at practices which alliances intend to maintain.

6.4.1 Practices to maintain



Two alliances do not intend to make any modifications to their analysis processes, two others remarked on their satisfaction with their categorisation. Mapping of outcomes to ToCs was mentioned by three alliances as having worked well. This has at least sometimes involved partners. In one alliance, the use of OH with MSC was again reported positively with MSC being useful for enabling substantiation of the significance.

6.4.2 Practices to change

Most evident among the responses is the interest among alliances in making analysis more participatory, more frequent, and doing more to support analysis and use.



Most frequently mentioned (by eight alliances), was the aspiration to have **more participation in analysis**, such as participation beyond MEL teams, the involvement of more levels in an organisation

or alliance, involvement of partners and other CSOs, and involvement beneficiaries or other external stakeholders in sensemaking. Interestingly, one alliance reported positively on their experience with involving wider stakeholders in a combined harvesting / sensemaking setting. Such a workshop potentially responds to the second main theme among the 'do differently' ideas: **make analysis more ongoing**, including by linking it closely with harvesting. Some six alliances identified this as an objective, with ideas including adding analysis categories to data collection software and the rating of significance and contribution by those describing outcomes.



Equally frequent among the comments were those identifying the need for **more support / space for analysis and sensemaking** and, in one case, a more systematic process. Others see the potential for wider use of existing data, including among country teams but noted the challenge that not all alliance organisations are equally open to adaptive management, hence the potential for programme steering informed by outcomes is not uniform.

Although no other major themes are evident in responses, several other challenges / possible improvements were identified indicating that analysis processes are likely to continue to develop if use of OH continues.

7 Assessment of OH experience

We turn now to look at the overall assessment alliances made about their OH experiences. **All but one of the respondents expect to continue to use OH beyond their D&D programme.** Clearly, there is great appreciation for OH. Alliances have also found considerable challenges. First, we look at the why alliances expect to continue using OH by assessing the advantages they identified and what those involved have appreciated. Next, we look at the main disadvantages, challenges with engaging participants and reasons why one alliance does not expect to continue with OH. Lastly, we consider observations on whether OH has worked better in some countries or contexts than others.

7.1 Advantages or why use will be continued

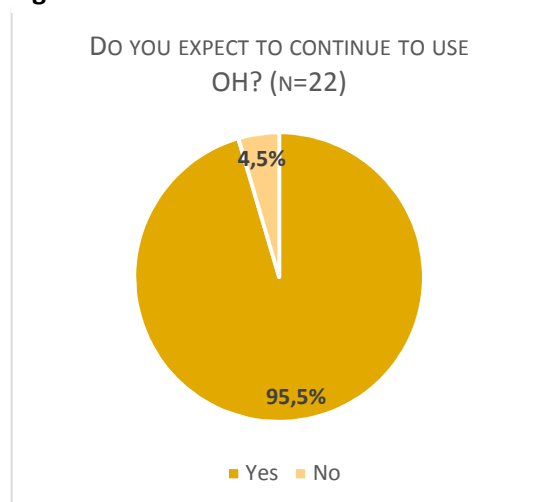
A key reason many plan to continue with OH is that there has been **wide ranging appreciation of OH** with several alliances noting that OH has been well received by partners, programme staff and, in one case, funders. One respondent remarked on the external context of their work being ever more complex, hence OH will continue to be useful.

For some, OH meets all PMEL needs such that they see no need for other approaches; others will continue to need other approaches for different types of programming or are curious to explore if other approaches, such as stories of change, can work better for e.g. the most local-level results.

Several attributes of OH are particularly valued. Let's start with the most widely valued of all: **the shift in focus from activities to changes.** This change of focus is highly appreciated and applies to programme staff, partners and other CSOs alike. One result is that: Partner CSOs have explicitly shown appreciation by telling us that OH helps them to deal more confidently with reporting requirements of other donors. There has been consistent feedback from partners that they have started using OH in their reporting to other donors.

Next we have **attention to unexpected and / or negative results**, something that 'opens up new thinking about the programme.' In one programme, 80% of outcomes have been unintended. Such outcomes can be invaluable for identifying which strategies to use.

Figure 12



[OH has been] 'well received by all involved and especially partner organizations/offices.'
Conducting Environments for Effective Policy

'[Outcome harvests are] possibly the most used data source by the Hague office & country offices collectively and to link the HQ/county office more closely.'
Towards a Worldwide Influencing Network

[shifting the focus from activities to changes] 'is important capacity building in itself and brought very clear information about what we achieved.'
Watershed

Equally prominent among responses was the **participatory nature of OH** which can strengthen links between PMEL, programme staff and partner CSOs and in turn build understanding of what an outcome is. Participation in harvesting can motivate implementation teams when even small successes are identified and celebrated. Above all, participation in OH is a process for ‘meaningful engagement on results with partners’, such that there can be reflection, interpretation and strategic thinking together. While participation in social change programming is highly valued in principle, so often it is elusive in practice. The experience of the alliances suggests OH is an approach that can help make genuine participation a reality through harvesting, analysis and sensemaking.

The usefulness of **OH alongside ToC or otherwise to inform adaptive management** is highly valued. The insights on programme contribution (and if interested, the contribution of others), the ability to track and learn from change processes to appreciate how small changes build into larger results and the learning about what strategies are and are not yielding results, all are valuable when implementing advocacy initiatives that by their nature are full of uncertainties that necessitate ongoing adaptation in implementation.

Less widely remarked upon but still notable was appreciation for **the flexibility of OH**. This has at least three dimensions: you are free to collect the information that is needed, in addition to the core outcome, significance and contribution; partners are not constrained by indicators but can describe changes that are actually meaningful to them; all can take an ‘open look’ at the programme, not just measure against expected results but be observant of what has actually been achieved, and how.

Noted by some was **the simple, intuitive nature of OH**. Telling a story fits with influencing and is a natural way for many to document change. The process is, essentially, easy to explain, even if the detail and precision needed for high quality data is an exacting process. In contrast, as we shall see below, others commented that actual use of OH is not so simple and demands particular skills.

Lastly, an idea that resonates with my experience is that **the value of OH is multi-dimensional**. This is implicitly reflected in the diversity of appreciation expressed by respondents and is specifically addressed in the following comment:



The potential [is] to use one methodology for multiple purposes (reporting, learning, improving programme quality, etc.). Linked to this: the fact that investing in capacity to do OH indirectly contributes to capacity on the programme's approaches, as it stimulates reflection on the real-life links between activity, output and outcome. Rutgers / Right Here Right Now.

7.2 Disadvantages and difficulties or why use won't be continued

Those that were positive about continuing their use of OH also indicated several disadvantages and difficulties in doing so. Before turning to these, let's consider the one alliance that has decided not to continue its use of OH. In short, it has found OH to be ‘too complicated’, as a result, they say, of not investing sufficiently in training and support. Considering that training and support / coaching are typically needed to successfully use OH, the decision not to continue is unsurprising.

While positive about continuing, the other alliances share several areas of concern or challenge.

‘The most suitable and meaningful [M&E approach] for advocacy initiatives that we have used as an organization.’
Rutgers / Right Here Right Now

‘It really provides information that we didn't have available before’
Oxfam Novib / Towards a Worldwide Influencing Network



Probably the most widely shared difficulty is the **significant time investment needed to use OH** and, linked to this, the **trade-off to be made between quality and quantity**. Reviewing (ping pong) in particular not only takes time but is a challenge to facilitate as it can be draining and depends on trust and an openness of those describing outcomes to accept feedback. The trade-off between quality and quantity can result in harvests missing outcomes and may be felt particularly acutely when there is time pressure to produce outcomes for reporting. One respondent commented emphatically: *[OH is] not to be used when the number of outcomes is very high (although grouping of similar outcomes is possible)* – Oxfam Novib / Towards a Worldwide Influencing Network. On the other hand, others find a balance by differentiating outcomes harvested for learning versus those for accountability.

‘Accountability comes at the second place. It is important, but I don’t mind losing some quality in the outcome reporting for the benefit of learning and reflection.’
Cordaid / Capacitating Change: Empowering people in fragile contexts

The second disadvantage or difficulty noted by respondents is related to the first: **capacity constraints**. This has several dimensions, including the need for initial training; coaching the identification of outcomes and writing of outcome statements; fostering an understanding of core OH concepts; the importance of follow-up training or coaching as new staff replace old; the great variety in peoples’ aptitudes such that some may never be good at writing concise, accurate statements, in which case processes are needed to provide support from others; sometimes limited PMEL capacity of partners and local offices such that they cannot effectively support the process internally; and the strong facilitation skills needed, but not always present, at each level.

‘...it is not an easy concept and it requires high investment in training, guiding, coaching.’
SNV Netherlands / Voice for Change

A difficulty related to the need for good facilitation skills is that of **participation**, specifically getting everyone involved who needs to be involved or, as one respondent put it, the challenge is to get *the “right people” involved in collecting and analysing the outcomes* - PAX / Freedom From Fear.

A fourth difficulty can come with **ensuring OH complements rather than competes** with other PMEL and reporting processes, such as indicator reporting and donor reporting.¹⁵

‘If OH is not embedded in ways of working with partners, it will remain a reporting tool.’
PAX / Freedom From Fear

When OH is introduced as an additional MEL process, its success is helped when a) the questions it answers complements what is learned from other PMEL approaches an organisation uses, b) OH data collection and reflection uses the same ‘rhythms and spaces’ as other monitoring processes and c) OH is valued as a learning process beyond accountability.

‘It has been hard to get to the point of really using the information for learning and programme steering at the country level’
Rutgers / Right Here Right Now

Last of the most commonly mentioned difficulties is **substantiation and validation**. Concerns included the time and effort needed, uncertainties

¹⁵ ‘Rhythms and spaces’ mentioned in the text box refers to the spaces and frequency an organisation has for discussion and decision making, as described in: Guijt, I. (2008) *Seeking Surprise: Rethinking Monitoring for Collective Learning in Rural Resource Management*, Wageningen, The Netherlands: Wageningen University

about how to do it, the perceived difficulty of it and the particular challenge of substantiation in sensitive contexts.

Other difficulties or disadvantages mentioned by two-three respondents are:

- Information overload both in number of outcomes and content of statements
- Missing information on the significance or meaning of an outcome or its effect on peoples' lives
- How to include community / stakeholder perspectives
- Lack of outcomes / outcome analysis in narrative reports submitted to headquarters
- The possibility of subjectivity or bias in outcome statements
- Harvesting of only positive outcomes / not documenting and learning from what is not changing
- OH cannot provide all that is needed on e.g. capacity development

7.3 Experiences using OH in different contexts

Of the 20 respondents providing comments, five had observed, experienced or anticipated differences in the feasibility of using OH in different countries. Reasons given include the challenge of addressing sensitive topics, such as LGBTI issues in the MENA region, that some actors feel uncomfortable speaking about, or limitations to M&E facilitation capacity in some locations.

From my experience, challenges can be overcome if it is possible to provide the required support and there is flexibility to select appropriate engagement and facilitation approaches.

Cultural and language differences were noted as factors determining the feasibility of using OH. Writing is not a preferred means of communication in all contexts. The outcome concept is not always well understood in some languages / cultures. Use of the active voice and affirmative language as recommended in OH goes against common practices in some languages / cultures.

Most commonly mentioned (seven respondents) was that ability to use OH is dependent on individuals, specifically the existence of a 'champion' to promote and coordinate the process and the facilitation skills. Crucially, the capacity to write using precise, concise English can vary enormously. It is not only the ability to write in such a way that varies, but also the motivation to do so: Perhaps surprisingly, the Dutch project officers are more reluctant to do the outcome harvesting writing. The partner organisations are very engaged in that.

8 Learning event discussions

Throughout the report, I identified topics that could usefully be discussed among D&D alliances and potentially more widely. My selection considered: a) where there was particular diversity of experience and hence learning opportunity among the alliances, and b) issues or challenges that are common among those using OH. These possible discussion topics are summarised in **Annex B**. The OH task force and survey participants selected three of these topics for exploration in two virtual ‘learning events’, the first on 8 July 2020 (28 participants registered), the second on 16 July 2020 (24 participants registered). The intention of the events was to further explore the topics through the sharing of perspectives and experiences, not to reach conclusions about how any one issue should be resolved going forwards.

Discussion topics

1. Outcome statement quality: what is 'good enough' quality and how can it vary; good practices and challenges at maximising data quality.
2. From use for reporting to use for learning.
3. Sustaining use of OH in organisations.

8.1 Discussion 1: OH data quality

Introduction to the discussion topic

The survey found that, on the one hand, alliances have found the harvesting of outcomes of sufficient quality to be a demanding process, whilst on the other hand the experiences of some alliances suggested there is potential for sharing of good practices in efficiently harvesting outcome statements of sufficient quality.

Central to any consideration of outcome quality is **OH Principle IX: Rigorously aim for credible enough outcomes**¹⁶:

Apply thorough evaluative thinking to ensure that the quality of the data and of the Outcome Harvesting exercise are as trustworthy as necessary for the primary users' principal uses.

Considering there are a range of possible principal uses of an OH exercise and data, there is inherently potential for variability in what will be ‘trustworthy as necessary’ and ‘credible enough’ in any particular application or use of OH. Indeed, some alliances found outcome statements with known weaknesses were still credible-enough to be used internally in an organization for reflection and / or programme steering, while a higher standard of quality and hence credibility was needed when using outcome data externally for reporting or evaluation:

‘Accountability comes at the second place. It is important, but I don't mind losing some quality in the outcome reporting for the benefit of learning and reflection.’ Cordaid / Capacitating Change: Empowering people in fragile contexts

‘two-thirds of outcome statements were SMART enough, [but] all were useful for sensemaking.’ IRC / Watershed

On pages 88-89 of Ricardo Wilson-Grau’s book on OH, he explores what makes OH data and the OH process credible¹⁷, or more specifically, credible enough for intended uses. In practice, there are a range of measures, or actions, that can be used across the harvesting process to enhance data credibility. The measures can be grouped as follows:

¹⁶ Wilson-Grau, R, 2019, page 183, Outcome Harvesting Principles, Steps and Evaluation Applications, IAP, Charlotte, NC.

¹⁷ Credibility is the quality of being believed or accepted as true. Wilson-Grau, R, *ibid*, page 88.

1. Outcome sources:
 - a) Agreement on who identifies outcomes and formulates outcome statements. Primary users agree with the evaluator / harvest coordinator on reliable primary and secondary sources.
 - b) The primary sources identifying outcomes and formulating outcome statements are the people who are most knowledgeable about the outcomes.
 - c) Primary sources are two or more people who mutually agree on the outcome statements, thereby creating data that reflects collective not individual knowledge. A variation on this is when outcome statements are internally verified i.e. their accuracy is confirmed by one or more people internal to the intervention whose outcomes are being harvested but not involved in formulating the outcome statements.
 - d) Secondary source material that you review to identify possible outcomes is written by someone other than the primary sources you engage with to review and finalise outcome statements.
 - e) Those describing outcomes go on record / the outcomes they describe are attributable to them.

2. Formulation of SMART outcome statements. The OH adaptation of SMART criteria is probably the most commonly used quality criteria in OH. Outcomes that are SMARTly described are, among other qualities, observable and hence verifiable, an absolutely key requirement in OH.

3. Supporting evidence that confirms the accuracy of an outcome and how an intervention contributed to an outcome. This has three components:
 - a) Sources are informed that the accuracy of their outcome statements may be tested with others, checked against supporting evidence, or through direct observation.
 - b) Sources provide references to supporting evidence and / or identify substantiators.
 - c) Harvesters verify the accuracy of outcomes by consulting supporting evidence and / or substantiators, or through direct observation.

The OH book does not tell us if any or all of these measures are essential, nor if any are more important than others, instead directing us to exercise our judgement by using 'thorough evaluative thinking' to determine what will make an outcome credible in light of the principal uses of principal users.

It is clear that some of the possible measures are optional as they would not be possible in some harvests. For instance, the use of secondary sources written by people other than primary sources is impossible when no such secondary material exists. More broadly, experience shows that other measures are optional as some users will require them, while others will not. The need for supporting evidence is an example of a measure that in practice some users require to consider an outcome credible enough whilst others do not.

Probably the most widely used measure is the requirement that outcome statements are SMART. Yet judgement is involved here too: does an outcome statement need to be fully intelligible to an outside reader if it is only to be used internally in an organisation? Or, to consider another common dilemma, does the contribution description need to contain all the information needed to determine plausibility if those using the outcome data know the contribution details and believe its influence on the social actor to plausibly have helped bring about the outcome? In contrast, when using an outcome externally, a higher level of detail or quality is needed because external outcome users a) lack the insiders' knowledge needed to fill information gaps in an outcome statement, b) generally

require a higher level of confidence in the accuracy of the data; often wants to see objective, measurable information; needs to be able to judge that the contribution is plausible; wants to see how an outcome is relevant to desired effect(s) of the intervention; and often are only interested in outcomes achieved during a specific timeframe.

Ultimately, it is the decision of users, guided by the harvester(s), to determine how to apply the SMART criteria, what other credibility measures are needed and how they should be applied to make outcomes credible enough for intended uses. **With the above considerations in mind, participants were invited to discuss:**

1. Attitudes to, and experiences with, the use of different quality criteria when using outcome data internally and externally.
2. Attitudes to, and experiences with, the use of supporting evidence to consider outcome data to be credible enough.
3. Good practices for harvesting credible enough outcomes.
4. Challenges when striving to harvest credible enough outcomes.

Discussion summary

Finding the best process for achieving required outcome quality has been an ongoing challenge for many if not most alliances. One discussion group summed up the challenge particularly well as *a constant act of balancing between what you want and what you can have*.

1. Attitudes to, and experiences with, the use of different quality criteria when using outcome data internally and externally.

On the question of whether the quality of internally used outcomes can differ (be less than) outcomes that are also used externally, experience has been mixed. Several discussants noted that initially, outcomes have been harvested for internal uses. Only later are some outcomes selected for reporting or other external communications. The question therefore is whether outcomes that were judged useful for internal uses were all credible enough for external uses or not.

Although a quantitative assessment of alliance experiences was not made, it is clear from the discussions that a majority of alliances have not differentiated the quality of outcomes used internally from those also used externally.

In contrast, for several alliances, a distinction in outcome statement quality has been made, not at the time of harvesting and use of outcomes internally but when selecting outcomes to use externally for reporting or when preparing a set of outcomes for an external evaluation. For at least three alliances, when outcomes were to be used externally, they were reviewed and those considered to be of insufficient quality were either removed from the dataset (if judged to be of low importance) or improved by going back to those who described the outcome for more information or clarifications.

No explicit or formal criteria for guiding the selection / rejection of outcomes or the improvement of outcome statements were mentioned by discussants. Instead, improvements were guided by experience or 'common sense', such that *'the specificity of outcomes was more carefully considered'*. Some types of weaknesses of outcomes were very concrete (e.g. a missing or incomplete date), in other cases it was a matter of judgement e.g. the plausibility of the contribution.

Groups summed up the quality distinction between outcomes good enough for internal versus external uses as follows:

Internally the S in the SMART could be less emphasized. For external [uses], detail and [the S in] SMART is important.

Internal: can you learn from it? External: is it verifiable?

In principle, differentiating outcome quality for internal and external uses has the potential to make it more efficient to harvest credible enough outcomes for internal uses, whilst allowing users to develop higher quality outcome statements for a selection of outcomes to be used externally. But there are risks, including:

1. By the time outcomes are selected for external uses, it may be more difficult, and in some circumstances impossible, to rectify the deficiencies in outcome statements. As some alliances found, attempts to resume ping pong after a reporting deadline can be challenging as the immediate requirement for outcomes has passed. Not resolving open questions on outcomes in a timely manner also risks people forgetting, losing motivation or moving position.
2. If the development of fully SMART outcome statements is left until an evaluation, the task of improving statements may be too large to accomplish with the evaluation resources and time available.

2. Attitudes to, and experiences with, the use of supporting evidence to consider outcome data to be credible enough.

Alliances have been widely varied in their requirements for supporting evidence. Although some noted that verification of outcomes with supporting evidence was not required by the MFA, many alliances have sought supporting information from those describing outcomes. For some, reference to supporting evidence was necessary to consider an outcome credible enough for any use. For more alliances, evidence, or reference to evidence, was routinely asked for but was not a make-or-break criteria for accepting an outcome statement as credible.

Less clear is the extent to which evidence was actually examined, though some commented that in practice checking evidence systematically was not feasible. Others streamlined the process by only verifying higher level / institutional / policy outcomes, not individual level outcomes.

Regardless of what the supporting evidence showed or which outcomes were verified, one alliance noted that the quality and hence credibility of outcome statements improved when partners were asked for evidence.

Asking for evidence / substantiators to be identified and informing sources that the accuracy of their outcomes may be investigated arguably goes a long way to enhancing outcome quality and therefore credibility, even if the accuracy of only a sample of outcomes is actually investigated.

3. Good practices for harvesting ‘credible enough’ outcomes

Foster OH thinking

Several discussion participants referred to the importance of partners / project implementers appreciating the OH way of thinking, not only because this motivates their participation in the OH process but also because when they understand the link between outcomes and their day-to-day work then outcome statement quality improves. Developing a clear understanding of ‘OH thinking’ can take time, but some found that reflecting on theories of change using outcomes can help build appreciation for OH. Another strategy noted by one alliance was to organise OH processes locally to build ownership. The goal, as one participant put it, is to build appreciation that OH is more than a technical approach, in other words it is a way of understanding that social change happens by influencing others to change behaviours.

How to start a harvest

A tip shared by several participants was to start a harvest relatively ‘light’ by identifying possible outcomes, referred to by one person as ‘baby outcomes’. The more promising or clearer baby outcomes are then developed into statements. A variation suggested by one organisation that can be used in a writeshop is to start with describing outcome ‘titles’ – such as one sentence ‘headlines’ – then to start developing statements from the clearest titles. Even if not all baby outcomes or headlines can immediately be developed into outcome statements because of the need to prioritise ping pong efforts, there is value in retaining the information as outcomes may materialise at a later date.

Brainstorming possible outcomes is often helpful for distinguishing outcomes from outputs (for those new to OH) and for rapidly capturing ideas about possible outcomes from all those describing outcomes without getting immediately stuck in the details of a small number of outcomes.

Another participant suggested to focus first on describing the outcome as an observable, verifiable behaviour change, then to move on to describing significance and contribution. The rationale here is that without an observable outcome description, there cannot be a SMART outcome statement.

There is merit in first describing the outcome but it is also prudent before investing time in formulating the outcome statement to check the intervention did indeed contribute to the behaviour change and that the change is indeed significant according to the criteria being used.

Limiting the number of outcomes

Recognition that formulating outcome statements is demanding and time consuming is one of the rationales behind **OH Principle IV is: Strive for less because it will be more useful**¹⁸. It follows that limiting the number of outcomes can help increase the quality of outcome statements. Several alliances recognise there is a balance to be struck between the number of outcome statements formulated and their quality. Several have also found it challenging to limit the number of outcomes formulated as statements. Thankfully, some alliances were able to share insights into how they had done so:

- Log possible outcomes then select some to formulate as statements every six months;

¹⁸ Wilson-Grau, R, 2019, pages 159-162, Outcome Harvesting Principles, Steps and Evaluation Applications, IAP.

- Select three-five outcomes for possible external use and obtain supporting evidence for these;
- Combine similar outcomes from different districts into a single outcome statement;
- Combine smaller outcomes into the contribution description of a single outcome statement;
- Use emotion instead of criteria to identify the most important outcomes: ask partners ‘what are you most proud of?’

Ping pong

The survey found that for many, the ping pong / review process was very challenging. However, some discussion participants remarked that ping pong can also have upsides: feedback can be much appreciated as it gives a feeling of recognition for the work done, which you don’t get if you just thank people for sending the reported outcomes. In one case, when feedback was provided by a Netherlands NGO PMEL, the interest shown was stimulating at the country level. Another positive experience shared by one alliance was the use of peer review between organisations. This allowed organisations to gain insight into the OH processes used by others.

When ping pong supports the work of those describing outcomes by clarifying thinking on what strategies yield results, provides insights into what has been achieved and provides outcome statements that together inform learning or communication, the process can be highly motivating. The challenge is often to find the best way to engage with each person.

Workshops / writeshops

Reinforcing the survey findings, several participants in discussions spoke of their preference for harvesting outcomes in a writeshop setting. Two advantages were noted: 1. the writeshop setting can be used to filter out outcome statements that are not SMART; and 2. This type of setting is better than email for getting SMART outcome statements.

4. Challenges when striving to harvest ‘credible enough’ outcomes

Echoing the survey findings, **building and sustaining capacity** was the most commonly referred to challenge with applying quality criteria / achieving sufficient data quality. The time taken for OH thinking and practice to become embedded, together with movement of staff, means that there is often an ongoing need to provide training / coaching / support materials to programme and PMEL teams. One suggested solution is to make OH part of staff induction.

One participant noted the different perceptions of data quality can create conflict: a tension exists between the story telling nature of OH (project team interest) and SMART data (M&E team interest).

Combining the harvesting of short outcome statements with the development of longer versions of statements or combining outcomes into an ‘outcome story’ can produce rich narratives akin to stories in MSC but with the precision of OH.

Harmonising interpretation of criteria or rubrics in harvests distributed across countries / organisations / harvesters is a challenge noted by some alliances. For instance, for one, significance was often interpreted differently: some people interpreted it as being in relationship to the final goal – is it significant in terms to what you want to achieve (“this is significant because it contributes to outcome 2”) – others noted only the significance in terms of the context of the outcome.

Care needs to be taken to ensure the description of significance is consistent. Similarly, as in any qualitative analysis, steps need to be taken to harmonise the classification of OH data between those participating in analysis in different teams, locations or over time.

8.2 Discussion 2: From use for reporting to use for learning

As many survey respondents reported using harvested outcomes for learning or programme steering / adapting as for reporting or accountability. Successful use for learning included using OH data and reflection for joint strategy setting, reflecting and adapting theories of change and for sharing successful strategies. Nonetheless, there was wide recognition among alliances that for many, OH has been used more, more successfully or more widely for reporting and somewhat less widely or successfully for learning. There was, in other words, considerable appetite for discussing how to strengthen the use of OH for learning, learning from good practices among some of the alliances. Survey responses indicated a particular interest in strengthening the use of OH for learning by:

- Increasing participation - getting the 'right people' involved in analysing outcomes
- Increasing the frequency of reflection / learning
- Providing more space and support for the learning process, including at the country level and by embedding OH reflection in work with partners

Who participates in analysis?

A key learning point identified by one discussion group was the need to consider the different levels of sensemaking needs and interests as well as outcomes collected. For instance, the outcomes collected that reflect individual behaviour changes might need different learning questions, different stakeholders and a different process compared to changes in behaviour on institutional level. And there will be different sense making needs at a local level which, such as including volunteers that have been involved in the implementation of the project, to sensemaking on a more strategic level by reflecting on the ToC.

The actual experiences of alliances ranged between those that had undertaken analysis mainly locally, mainly or only at the headquarters level, or both. Local-level good practices are:

- Quarterly quality check and reflection to build ownership
- Providing a manual of options for learning questions and online reflection exercises, providing facilitation support where needed.
- Mapping outcomes to pathways of change.

At headquarters, the synthesis of data involves use of numerical summaries to indicate patterns but to check assumptions it is necessary to examine the content of outcome statements.

Time and space for analysis

Two steps in the analysis process were identified: 1. Organise the outcomes, create timeline, select most significant; 2. Discuss more deeply using learning questions.

While participation can be highly valued in OH analysis, it can take time to develop the trust needed for all participants to openly indicate their assessment of what outcomes mean, their significance or the project's contribution.

Uses of analysis

In addition to the widely practiced mapping of outcomes to theories of change to check for relevance, test assumptions and inform strategies, alliances have used analyses to address particular interests:

- Capacity analysis of the alliance using 5-C model¹⁹
- Focus on network analysis to examine alliance building
- Focus on coalition building

In one case, a participant reported that analysis of outcome data had informed a decision to stop working with a partner because of the lack of outcomes. This highlights the tension inherent when using OH for both accountability and learning simultaneously.

8.3 Discussion 3: Sustaining OH in organisations

The survey indicated a near universal optimism among about continued use of OH. Nonetheless, in light of the imminent end of the D&D programme there was considerable interest in discussing how to sustain OH in organisations.

One way of understanding what is needed to sustain OH in organisations is provided by a set of **'enabling factors'** that have proven to be helpful for determining if Outcome Mapping (OM), a sister approach to OH, is suitable for an organisation of intervention²⁰.

The D&D alliances and most likely other projects of the alliance member organisations, including but not limited to advocacy, indicate the presence of the first of the three essential enabling factors: the existence of complexity, defined here as being uncertain about the results you expect or having uncertainty about processes by which results will be achieved. Similarly, use of OH and participation in the survey suggest essential enabling factors #2 and #3 are also present across the 22 alliances participating in the survey. The minimum requirements for sustaining OH thinking and the essentials of OH practice at a relatively small scale are therefore currently in place. The presence of the essential enabling factors helps explain the survey finding that all but one alliance expects, despite notable challenges, to continue using OH after the D&D programme. Considering that the majority of alliance members had not used OH before the D&D programme, the optimism about sustaining OH use is striking and suggests that one or more 'optional enabling factors' are widely present.

Essential enabling factors

All three are required if OH / OM is to be used successfully / sustained:

1. Existence of complexity in the project / organisational context
2. Willingness to act on complexity by using OH/OM thinking to change practices
3. Champions and support, someone to celebrate and motivate others as they learn to use OH/OM

Optional enabling factors

One or more of the following will help sustain OH/OM:

1. Funder support
2. Executive level support
3. Learning culture
4. Wide appreciation of learning/results-oriented M&E
5. Sufficient resources

For use of OH to continue at or beyond the scale of use by the D&D alliances, one or more 'optional enabling factors' will be needed.

¹⁹ https://www.betterevaluation.org/en/resources/discussion-paper/capacity-change-performance-study_5c-approach

²⁰ Accessed 02.08.20: : [10 years of Outcome Mapping: Applications, Adaptations & Support | https://www.outcomemapping.ca/resource/10-years-of-outcome-mapping-applications-adaptations-support](https://www.outcomemapping.ca/resource/10-years-of-outcome-mapping-applications-adaptations-support)

Challenges to continuing OH use that discussion participants highlighted include:

- the dominance of indicator-based M&E in alliance member organisations,
- the need for sufficient budget to optimise participation in OH through workshops, and
- the high existing M&E demands on programme and PMEL staff because of the requirements of others donors.

One group noted the need to demonstrate the added value of OH, using the D&D experience as an example. The survey findings show there is much potential to do just this.

Notably, some discussion participants remarked on the importance of champions among the programme / project staff. Unexpected outcomes can help stimulate greater ownership when they are used as evidence to inform a shift in focus or strategy of a project. As project / programme managers become champions, not only does OH become demand driven from the field but coordination and support for the OH process can be shared beyond the PMEL staff.

The role of OH champions in both PMEL and project teams cannot be underestimated in sustaining OH. Demand for OH from project / programme staff is particularly useful. Even one project staff member using OH can create examples of OH practice that can then inspire others.

Annex A: Survey respondents

Alliance	Lead Organization	Contact
Advocacy for Change	Solidaridad	Victoria Graham <victoria@solidaridadnetwork.org>
Building Capacity for Sector Change	Rainforest Alliance	Tessa Witte-Laan <TWitte-Laan@ra.org>
Capacitating Change: Empowering people in fragile contexts	Cordaid	Koen Faber <Koen.Faber@cordaid.org>
Citizen Agency Consortium	Hivos	Karel Chambille <kchambille@hivos.org>
Civic Engagement (Convening and Convincing)	ICCO	Dieneke de Groot <d.degroot@icco.nl>
Conducive Environment for Effective Policy Influencing	NIMD	Nic van der Jagt <NicvanderJagt@nimd.org>
Every Voice Counts	Care Nederland	Mirjam Locadia <locadia@carenederland.org>
FFF - Freedom from Fear	PAX	Yvonne Es <es@paxforpeace.nl>
Girls Advocacy Alliance	Plan Nederland	Helen Evertsz <helen.evertsz@planinternational.nl>
Green Livelihoods Alliance	Milieudefensie	Maya Verlinden <maya.verlinden@milieudefensie.nl>
HSA Partnership (Health Systems Advocacy for Africa)	AMREF	Maarten Kuijpers <Maarten.Kuijpers@amref.nl>
N/A	RNW-Media	Elianne Anemaat <Elianne.Anemaat@rnw.org>
No News is Bad News	Free Press Unlimited	Michael Pavicic <pavicic@freepressunlimited.org> Tim Schoot Uiterkamp <schootuiterkamp@freepressunlimited.org>;
Partners for Resilience	Rode Kruis - Nederland	Conny Hoitink <Conny.Hoitink@wetlands.org>

Partnership for Rights, Inclusivity, Diversity and Equality (PRIDE)	COC Nederland	Mayanka Vij, M&E officer < MVij@coc.nl > Bram Langen < blangen@coc.nl >
PITCH (Beat the Aids Epidemic)	Aidsfonds	Maartje van der Meulen < maartjevandermeulen@aidsfonds.nl >
Prevent up Front	GPPAC	Paul Kosterink < p.kosterink@gppac.net > Wieteke Overbeek: < w.overbeek@gppac.net >
Right Here Right Now	Rutgers	Inge Vreeke < i.vreeke@rutgers.nl >
Shared Resources, Joint Solutions	IUCN	Marielouise Slettenhaar-Ket marielouise.slettenhaar-ket@iucn.nl Gunilla Kuperus < gkuperus@wwf.nl >
Towards a Worldwide Influencing Network	Oxfam Novib	Anne Oudes < Anne.Oudes@oxfamnovib.nl >
Voice for Change	SNV	Margriet Poel < mpoel@snvworld.org >
Watershed-Empowering Citizens	IRC	Conny Hoitink < Conny.Hoitink@wetlands.org >

Annex B: Themes for exchange and learning

The following themes and topics for possible further exchange and learning were identified through analysis of the survey findings. Three topics were discussed at two events to date: 1. Outcome quality; 2. From use for reporting to use for learning; 3. Sustaining OH use.

Themes	Discussion topics
Quality vs quantity vs effort trade offs	<p>‘Good enough’ for learning vs reporting: Several alliances struggled to consistently formulate SMART outcome statements yet some indicated even less than SMART outcome statements were still credible enough for some uses. Can an outcome be ‘good enough’ for learning and steering, even if not good enough for reporting? For alliances that have distinguished outcomes for internal use only from those that can be used externally, what different quality standards or requirements for supporting evidence have been used? Is triangulation needed for both? Is substantiation needed for both?</p> <p>Improving outcome quality: What aspects of the timing / frequency / use of peer review or other features of the harvest and review process are most useful for increasing outcome quality? How to develop and maintain strong engagement throughout ping pong?</p> <p>Limiting the number of outcomes harvested: When the number of draft outcomes is too great to develop all into SMART outcomes, how can the number of outcomes be made more manageable? Was ‘importance’ defined by alliances and was it possible to apply the importance criteria consistently to guide the inclusion / rejection of outcomes from a harvest?</p>
Harvesting ‘hard to reach outcomes’	<p>Some succeeded in harvesting unintended or negative outcomes while some did not. What good practices can be shared?</p> <p>Capacity change outcomes: some succeeded in harvesting outcomes that evidence capacity changes, some find OH doesn’t work for such results. Conversely, some succeed in capturing advocacy outcomes, some struggled. Why? What practices can be shared?</p>
From use of OH for reporting to its use for learning and programme steering	<p>Many want to strengthen analysis and sensemaking, not least participation in it. Many have succeeded. A sharing opportunity.</p> <p>How to link sensemaking to the harvesting process in practice e.g. in a workshop flow: practices to share.</p> <p>How to integrate analysis in data collection tools – examples to share.</p>
Participation and diversity of perspectives	<p>Participation: how to give voice to all those needed to make harvesting informative and useful. How to get all involved.</p> <p>Combining harvesting / analysis / sensemaking. Getting the balance right - the latter two being most motivating.</p>

	Experiences using MSC to add perspectives on significance in OH.
Substantiation	<ul style="list-style-type: none"> • Challenges noted include criteria to select a sample to be substantiated, how 'external' should substantiators be, how often should substantiation be done and when. • Substantiation for accuracy, enriching understanding of outcome and / or contribution AND different perspectives on significance. Discuss. • Often substantiation was led by staff. Was this out of choice or because of budget constraints? Any implications for credibility of the process? • Outcomes of advocacy and influencing are often sensitive. Explore options for substantiating advocacy outcomes / in sensitive contexts • Consider the proportion of positive substantiator responses needed to confirm an outcome as substantiated • How should the accuracy / credibility of unsubstantiated outcomes be interpreted if substantiation thresholds are not used? • Contrast triangulation / verification / substantiation in OH with other qualitative evaluation approaches used by the alliances • What is the actual experience of using OH data in terms of users challenging its credibility?
Evaluation use of outcomes harvested by monitoring	Can investment in monitoring be partly justified by its value for external evaluations? Has the quality of outcomes and associated analysis been sufficient for external evaluations? How satisfied are alliances with the use of monitoring data by evaluators: are evaluations better in quality or scope, are they better value as less data collection is needed? To what extent has the evaluation use of OH data been limited when negative or unintended outcomes are missing?
Sustaining OH in organisations	<ul style="list-style-type: none"> • Is OH easy / simple vs or not an easy concept and need for high investment? • Challenge of maintaining capacity / refresh capacity with changing staff • Integration of OH in funding proposals – tips to share?