Indicator Aggregation in International Cooperation

How to Summarise Results at Programme Level

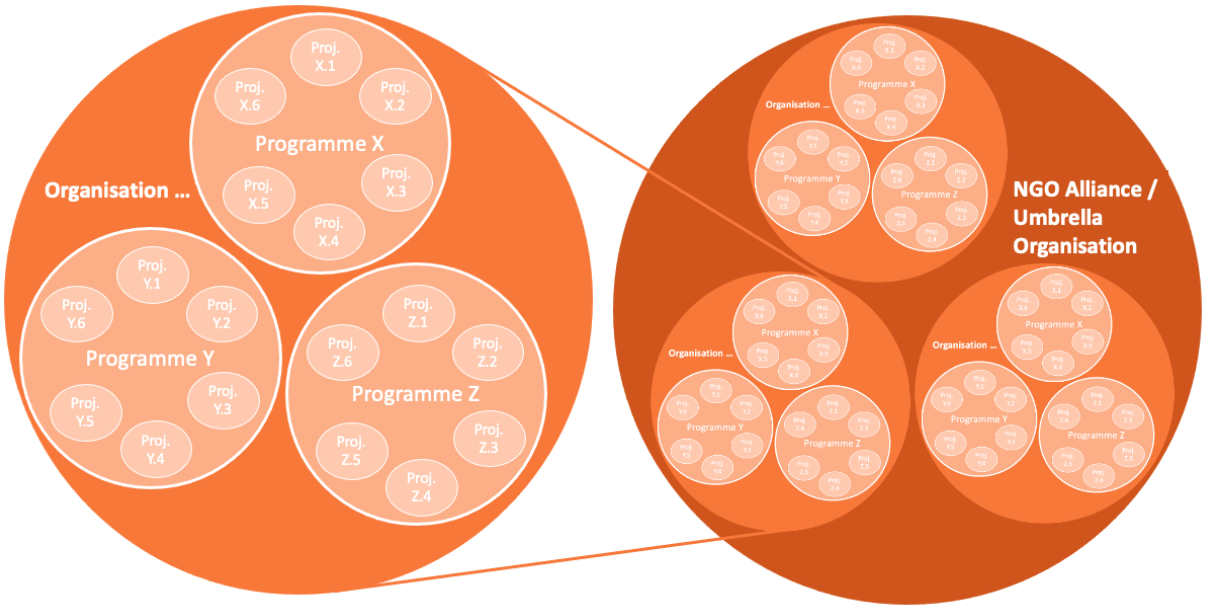
Measuring results and change is mainstream for development and humanitarian projects. In recent years, doing so has become increasingly common at programme or organisational level as well. This facilitates organisational steering, supports strategic processes, and helps to ensure accountability to project participants, donors and the wider public. However, summarising project achievements at higher levels is challenging. Results from multiple projects at diverse stages and of different durations have to be aggregated. Such M&E systems have been referred to as "complex", contrasting simpler project systems.[1]

This paper sheds light on key factors worth considering when aggregating results in complex M&E systems[2]. Hopefully, this will facilitate smoothly running M&E systems and set free untapped potential to steer programmes and organisations towards more effectiveness and impact.

Aggregating results - at what level?

At what level do you want to be accountable, and to whom? What levels are important for strategic decision making? Try to keep your system as simple as possible. As a rule of thumb, reduce the levels of aggregation to a minimum and let the decision be guided by how your organisation will use the results. Simplicity is king/queen when it comes to complex M&E systems. On the next page is a non-exhaustive list of different levels information can be summarised at.

> ***… reduce the levels of aggregation to a minimum and let the decision be guided by how your organisation will use the results.***



[1] Simister, N. (2019). *Complex M&E Systems: Raising standards, lowering the bar* (No. 6; Praxis Series Paper), p.5. This paper complements the cited work by specifically showing how to go about with indicator aggregation.

[2] "Complex M&E system" is used interchangeably with "programme" or "organisational M&E system".

Level	*Description*
Country programme	Country programmes may be important administrative entities for management and strategic steering. Organizations may choose to summarise information for each country.
Thematic programme	In case thematic programmes are important, organisations may decide to summarise information thematically as well.
Regional programme	Some organisations are structured in regional clusters, and hence it could make sense to summarise information at regional level.
International Cooperation Programme / Organisational Strategy	Organisations often label their development cooperation and humanitarian aid strategy as "International Programme". This usually shows information at the level of the whole organisation in the area of international cooperation, or large parts of it.
Multi-organisation Programme / Strategy	In some cases, organisations join forces to come up with common programmes / strategies. Hence, an additional level is added, and summarisation happens across organisations (ref. above illustration for NGO Alliances or Umbrella Organisations).

Aggregating results - which ones to choose?

You will not be able to aggregate all your project results. Make choices. As a rule of thumb, focus on very few outcome indicators and a small number of output indicators. Output indicators are often easier and less resource intense to measure and aggregate. Answering the following questions may help to make your choice:

You will not be able to aggregate all your project results. Make choices.

- ☐ Does the indicator show strategically relevant information for your programme? Does it highlight an important area of work?
- ☐ Do the indicators you choose represent a significant part of your programme?
- ☐ Will the information be useful at project level as well?
- ☐ Can the indicator be measured among a portfolio of projects? Is it simple enough?
- ☐ What additional resources do projects require to measure it? How can these additional resources be kept low?

Indicator types

The following types of quantitative as well as qualitative indicators are most relevant for showcasing programme level results.[3] In the remainder we will focus on aggregate and basket indicators.

Direct indicators:
- Centrally and directly collected at programme level.
- Do not rely on project information.
- Feasible only in few circumstances: big organisation with plausible contribution, possibility to conduct centrally led global survey, secondary data available (ref. cited work). If feasible, the best option to showcase results.

Aggregate indicators:
- Indicators are collected at project level in the same (standardised) manner (i.e., definition, method, tools).
- Collected values must be from projects in similar stages and refer to comparable timescales.
- Feasible for number and percentage indicators.
- Project values are aggregated to programme level.

Basket, bucket, or framing indicators:
- Defined broadly to identify an area of interest at programme level.
- Cannot be directly collected.
- More specific indicators are collected at project level and aggregated into a broader "basket".
- Feasible for number indicator and qualitative indicators.
- Difficult to define a basket for percentage indicators.

Qualitative indicators / information:
- Qualitative indicators and information are best summarised in a narrative form with a structured analysis method. They can also come in the form of basked indicators.
- If the information is large, a text analysis tool (e.g. NVivo, ATLAS.ti) or Large Language Model (LLM, e.g. ChatGPT) can be used to help with summarising.

[3] Simister, N. (2019). *Complex M&E Systems: Raising standards, lowering the bar* (No. 6; Praxis Series Paper), p.32.

How to treat number indicators?

Outcome and output indicators follow different logics.[4] Outcome indicators usually try measuring the "state" or "situation" within a target population and can usually not be fully attributed to a single intervention. They are often formulated as a percentage, but in some cases can also be formulated as a number.[5] Output indicators, on the other hand, count or "quantify" the output generated by the main activities. Most often output indicators are formulated as a mere number. The aggregation of number indicators is straightforward – different values of projects contributing to a programme can simply be summed up. The following tables illustrate aggregation for basket and aggregate indicators.

The indicators you aggregate must consistently follow either incremental or cumulative logics.

The indicators you aggregate must consistently follow either incremental or cumulative logics. Cumulative indicators measure what is accomplished from the start of the project up to the reporting date. Incremental indicators, on the other hand, measure the accomplishment during the reporting period only. The indicator name should be clear about whether the data is incremental or cumulative.[6] Additional complexity is added if you want to count "unique persons". Here again, it is crucial that the data you aggregate is consistent. To keep M&E systems as simple as possible and ensure they generate meaningful data, we do suggest remaining with either incremental or cumulative logic for most, if not all, of your indicators. If counting unique project participants is not necessary, it is best avoided, as it often becomes unmanageable.

Basket indicator to measure Output

	Level	*Indicator*	*Y1*	*Y2*	*Y3*
Incremental	Project 1	# of persons trained on improved business skills	1'000	800	1'200
	Project 2	# of persons trained in quality health services	2'000	3'000	0
	Project 3	# of persons trained in agro-ecological practices	11'000	10'000	8'000
	IC Programme	Basket indicator: # of persons trained on skills development	14'000	13'800	9'200

	Level	*Indicator*	*Y1*	*Y2*	*Y3*
Cumulative	Project 1	# of persons trained on improved business skills since the beginning of the project	1'000	1'800	3'000
	Project 2	# of persons trained in quality health services since the beginning of the project	2'000	5'000	5'000
	Project 3	# of persons trained in agro-ecological practices since the beginning of the project	11'000	21'000	29'000
	IC Programme	Basket indicator: # of persons trained on skills development since the beginning of the programme	14'000	27'800	37'000

[4] Interaction/Malacarne, N. (2023). *Indicator Monitoring in International Cooperation: Overcoming Common Challenges* (Issue June).
[5] Ibid. - refer for examples.
[6] "# of persons with access to safe drinking water" can be interpreted as cumulative. "# of persons using safe drinking water" is rather incremental. However, this is not entirely clear. Hence, it is advised to specify by using "cumulative" / "incremental" tags.

Cumulative, unique person[7]	*Level*	*Indicator*	*Y1*	*Y2*	*Y3*
	Project 1	# of persons trained on improved business skills for the first time since the beginning of the project	1'000	1'400	2'000
	Project 2	# of persons trained in quality health services for the first time since the beginning of the project	2'000	3'500	3'500
	Project 3	# of persons trained in agro-ecological practices for the first time since the beginning of the project	11'000	16'000	20'000
	IC Programme	Basket indicator: # of persons trained on skills development since the beginning of the programme	14'000	20'900	25'500

Aggregate indicator to measure Output

Incremental	*Level*	*Indicator*	*Y1*	*Y2*	*Y3*
	Project 1	Aggregate indicator: # of persons trained in agro-ecological practices	5'000	7'000	3'000
	Project 2		15'000	10'000	0
	Project 3		4'000	5'000	8'000
	IC Programme		24'000	22'000	11'000

Basket indicator to measure Outcome

Incremental	*Level*	*Indicator*	*Y1*	*Y2*	*Y3*
	Project 1	# of local policy changes improving the situation of Leprosy affected people	1	1	3
	Project 2	# of local policy changes improving the situation of the Adivasi indigenous community (Bangladesh)	0	0	2
	Project 3	# of local policy changes improving the situation of the Karamojong indigenous group (Uganda)	1	0	2
	IC Programme	Basket indicator: # of local policy changes improving the situation of marginalised and disadvantaged groups	2	1	7

How to treat percentage indicators?

This section requires basic knowledge of statistics. ⚠

Outcome indicators, on the other hand, are often formulated as percentage values. Percentage values show the relationship of two numbers:

$$\frac{\text{Persons fulfilling indicator criterion (numerator)}}{\text{Target population, usually the surveyed population (denominator)}} \text{ x } 100 = \% \text{ indicator value}$$

They are most interesting if compared over time. The trend over time will be your project achievement. This needs to be reflected at programme level. Percentage indicators usually follow a clearly defined methodology, such as representative household survey with standardised survey questions. We limit this section to the aggregate indicator type.[8] Percentages can be aggregated by using the simple average, median or to reach the most accurate

When aggregating you must give due consideration to the stage of a project and the timeframe of data collection.

[7] This example assumes that each year 50% of persons are re-trained, and hence not counted again (de-duplication).
[8] The cumulative/incremental differentiation is secondary for outcome indicators which are collected by representative household surveys.

results, the weighted average[9]. The median is a good choice, as it is simple and larger / lower numbers do not skew the aggregated value. When aggregating you must consider the stage[10] a project is in and the timeframe of data collection. This makes the aggregation of percentage indicators complex.

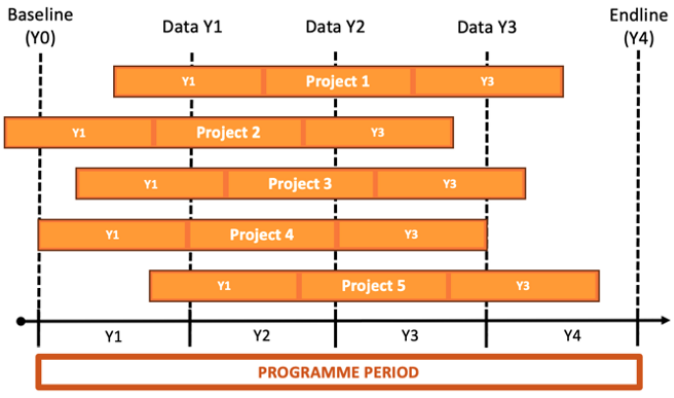
The aggregation in below table implies the programme and project durations are three years and overlap (correlate) fully. All projects are at the same stage, and the collected data refer to the same timeframe. This is crucial, because outcome indicators usually measure the state or situation within a target population over time. Hence, the same projects must be compared over time and set in relation with others at similar stages. In reality however, you will find situations where there is partial or no correlation between project and programme periods.

In reality, you will find situations with no or only partial correlation between project and programme periods.

Aggregation in an ideal situation

Level	*Indicator*	*Programme*			
		Baseline	*Y1*	*Y2*	*Endline*
Project 1	Aggregate indicator: % of targeted population experiencing moderate or severe food insecurity (FIES)	70%	75%	73%	71%
Project 2		63%	60%	61%	50%
Project 3		55%	45%	33%	26%
Project 4		80%	77%	63%	65%
Project 5		90%	75%	67%	40%
IC Programme	**Mean** (average) as aggregation rule	72%	66%	59%	50%
	Median as aggregation rule	70%	75%	63%	50%

In situation A, we extrapolate the project level data collections onto the programme time period: Projects 2, 3 and 4 are aggregated for the programme baseline. Even though project 3 starts in the second quarter of Y1, it may still make sense to use it for baseline calculation, assuming the baseline values have been collected in year 0. You want as many projects as meaningful to contribute to your baseline value. Projects 1 and 5 start only in the second half of year 1, which is why we decided not to include them in the programme baseline value. Project baselines are unlikely to arrive when the programme commences. For the programme's year 1 result, year 1 project values of 2, 3 and 4, as well as baselines of 1 and 5 are aggregated. For year 2, 3 and endline of the programme, all projects contribute information, although at different stages. Projects 2, 3 and 4 contribute the same information twice (in year 3 and



Situation A: Partial correlation between between programme and project durations.

[9] If you aggregate very small with large projects and/or in case you aggregate information from only few projects, it may make sense to use the weighted average to aggregate. You can weigh percentage values by their denominator (total target population, not just surveyed population!). But this aggregation method is associated with additional complexity. You will need to have numerator and denominator values reported separately and be aware of the targeted population. The targeted population may change over time which will make aggregation difficult to manage.

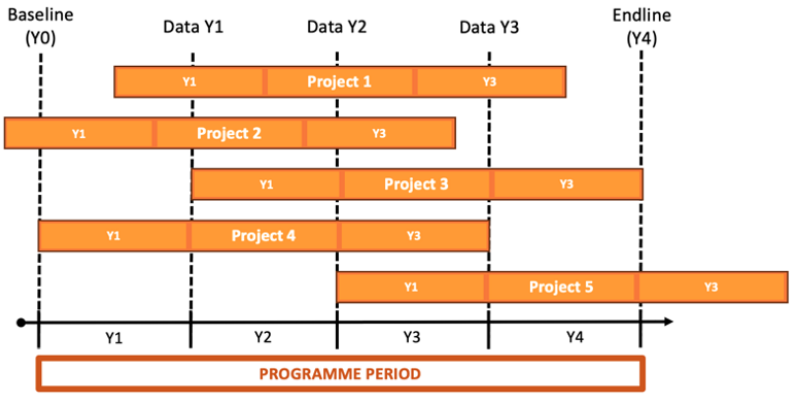
[10] By project stage (or phase) we mean, if a project is in its planning stage, inception stage, at mid-term, towards the end, or already completed.

endline). You can see that the median option leads to a higher level of food insecurity in years 1 and 3, while in the mean option food insecurity increases in year 1 only. It is important to explain in your reporting that part of the increase is due to additional projects contributing to the aggregated value.

Aggregation in situation A

Level	*Indicator*	*Programme*				
		Baseline[11]	*Y1*	*Y2*	*Y3*	*Endline*
Project 1	Aggregate indicator: % of targeted population experiencing moderate or severe food insecurity (FIES)		BL[12]: 70%	Y1: 75%	Y2: 73%	EL[13]: 71%
Project 2		BL: 63%	Y1: 60%	Y2: 61%	EL: 50%	EL: 50%
Project 3		BL: 55%	Y1: 45%	Y2: 33%	EL: 26%	EL: 26%
Project 4		BL: 80%	Y1: 77%	Y2: 63%	EL: 65%	EL: 65%
Project 5			BL: 90%	Y1: 75%	Y2: 67%	EL: 40%
IC Programme	**Mean** (average) as aggregation rule	66%	68%	61%	56%	50%
	Median as aggregation rule	63%	70%	63%	65%	50%

Situation B looks different. Project 5 will not be considered for the programme. For this illustration, we have decided that project and programme need more than two years overlap. Adding a new project baseline when the programme is already at mid-term would worsen the aggregated programme result (as occurred in situation A, Y1). Further, the project endline will not be available during the project duration. Project 5 will not be able to contribute the full range of change.



Situation B: No correlation between programme and project durations.

Aggregation in situation B

Level	*Indicator*	*Programme*				
		Baseline	*Y1*	*Y2*	*Y3*	*Endline*
Project 1	Aggregate indicator: % of targeted population experiencing moderate or severe food insecurity (FIES)		BL: 70%	Y1: 75%	Y2: 73%	EL: 71%
Project 2		BL: 63%	Y1: 60%	Y2: 61%	EL: 50%	EL: 50%
Project 3			BL: 55%	Y1: 45%	Y2: 33%	EL: 26%
Project 4		BL: 80%	Y1: 77%	Y2: 63%	EL: 65%	EL: 65%
Project 5		-	-	-	-	-
IC Programme	**Mean** (average) as aggregation rule	72%	66%	61%	55%	53%
	Median as aggregation rule	72%	65%	62%	58%	58%

[11] The descriptions in the first row, refer to programme baseline, Y1, Y2 or Y3 results.
[12] The descriptions in the cells, refer to project baseline, Y1, Y2 or Y3/endline results.
[13] Endline

There is no one rule on how to aggregate percentage aggregate indicators from project to programme level. An M&E team will need to manage a set of rules to ensure which values are aggregated to show programme results. The more projects you aggregate the more complex it will get to manage.

Other options

If you want to keep your system simple, it may be advisable to select only a few projects contributing to aggregate indicators – these can become M&E flagship projects.[14] This selection should include projects that correlate well with the programme duration and are exemplary for the programme. For example, projects 2 and 4 in situation B.

If you want to keep your system simple, it may be advisable to select a few projects contributing to aggregate indicators – these can become M&E flagship projects.

Another option is to generate a random (stratified) sample of people from within selected projects, and then infer results across their combined target populations. As in the situations outlined above, you would define criteria for projects to be included (i.e. projects with a similar starting date as the programme, projects not older than two years, etc.). Among included projects you can then randomly sample people and calculate your indicator across the combined sample. This is very similar to the situations outlined above where project values are aggregated using the weighted average, with the difference that measurement would be more centrally managed in this option.[15]

If you want to learn more about complex M&E systems, do go through the paper "Complex M&E Systems: Raising standards, lowering the bar".[16] It is an excellent piece to show how M&E can take place at programme level. This paper has built on some of the points, giving concrete examples of different programme indicator types and deepened the question of how to tackle issues with percentage aggregate indicators.

Qualitative information and indicators

Do keep in mind that qualitative indicators allow aggregation and summarisation of your project achievements too, as the following example illustrates.

Qualitative basket indicator

Level	*Indicator*	*Y1*[17]	*Y2*	*Y3*
Project 1	Evidence that CSOs influenced policy improving situation of Adivasi indigenous minority community (Bangladesh).	No evidence. (CSO trained on policy advocacy.)	CSOs submit a letter to the government urging to improve rights of Adivasi community.	One CSO participates in multi-stakeholder platform supporting policy implementation.

[14] Another option discussed while writing this paper is to show outcome indicators as mere numbers by summing up all the numerators of the percentage values. However, this has the disadvantage of losing the information of at which stage a project stands. Further, interpreting baseline values becomes difficult.
[15] This option is very close to the direct indicator type explained in the table on page 2.
[16] Simister, N. (2019). *Complex M&E Systems: Raising standards, lowering the bar* (No. 6; Praxis Series Paper).
[17] For brevity of this paper, the qualitative explanations are kept short. In reality, you may require more detailed information. You could define a few key dimensions / questions per indicator to report on in a structured manner.

Level	*Indicator*	*Y1*[17]	*Y2*	*Y3*
Project 2	Evidence that CSOs influenced policy improving situation of Karamojong indigenous minority group (Uganda).	No evidence. (CSO trained on policy advocacy.)	Three CSOs are invited by the government to attend a consultation conference on new policy.	Two CSOs invited to partake in a technical working group elaborating new policy.
Project 3	Evidence-based policy advocacy resources receive recognition.	CSO conducts research on minority group.	CSO publishes a policy brief on the discrimination of minority group.	The policy brief is presented at conference of the government upon invitation.
IC Programme	Basket indicator: Level of policy influence by supported CSOs (none, low, medium, high)	None: CSOs mainly trained, and policy research activities undertaken.	Low: CSOs speak up and publish policy brief. They are consulted by the government.	Medium: CSOs participate in policy formulation and implementation. Publications get recognition.

Moreover, beyond demonstrating change through pre-defined indicators, there are various qualitative M&E methods to capture intended and unintended programmatic changes (e.g., Most-significant change, Outcome Harvesting, etc.). Furthermore, reporting systems will also capture and aggregate narrative information. As qualitative methods are usually simpler to manage, they can lead to more useful information. Certainly, they help to triangulate your quantitative information.

Acknowledgement
Gratitude goes to the reviewers who provided meaningful input substantially improving the quality of this paper: Michael Angelopoulos, Mikaël Amsing, Nigel Simister, Ilona Paganoni, Tatiana Jaramillo, Thomas Gass.

Learn more about Interaction's international programme *Learning 360*:

- Programme Overview – German & French
- How we understand Impact – German & French

---

info@interaction-schweiz.ch; www.interaction-schweiz.ch.

MEL Paper Series. November 2023.
Author: Nicola Malacarne, Programme Coordinator / Advisor Monitoring, Evaluation & Learning (MEL), Interaction; nicola.malacarne@interaction-schweiz.ch.

Suggested citation: Interaction/Nicola Malacarne (November 2023). Indicator Aggregation in International Cooperation: How to Summarise Results at Programme Level. MEL Paper Series.