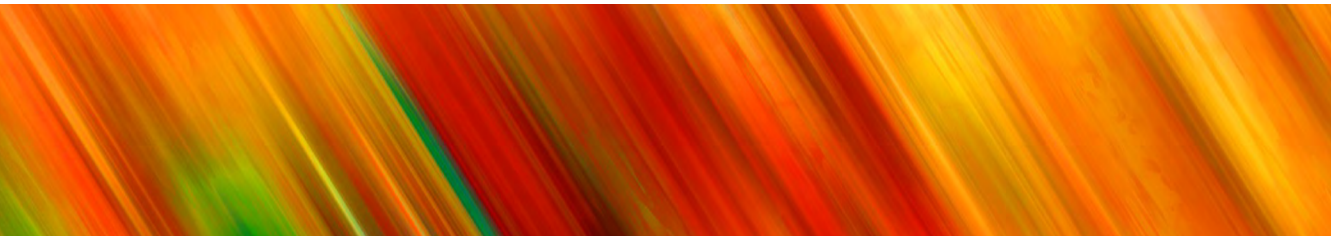
CMCFinland
Volume 11: 3/2023

THREE STEPS TO EVALUATING EFFECTIVENESS – LEARNING FROM FINNISH REPORTS ON AFGHANISTAN

JYRKI RUOHOMÄKI & JOHANNA HAKANEN



CMC Finland Working Papers analyse civilian contributions to peace operations and include recommendations for developing practical capabilities and improving effectiveness.

Author
Jyrki Ruohomäki
Head of Research
CMC Finland

Author
Johanna Hakanen
Research and Development Specialist
CMC Finland

Contact
research@cmcfinland.fi
Crisis Management Centre (CMC) Finland
Ratapihantie 9, 6th floor, 00520 Helsinki
P.O. Box 1325, FI-00521 Helsinki

ISSN 1797-1667

Layout and print: PunaMusta Oy

Content

1 Introduction

Participation in peacekeeping and later in crisis management has been a key element in Finnish foreign and security policy. During the Cold War, participation in UN peacekeeping missions, for example, provided Finland, and in particular the Finnish Defence Forces, with an opportunity to interact with Western colleagues, which would have been impossible otherwise. After the Cold War, Finland has considered both military and civilian crisis management missions and operations as avenues to bolster its international visibility and build partnerships. The government manifesto from 2019 sets the aim for Finnish foreign policy to be "larger than its size" internationally (VN 2019), putting Finland amongst the biggest nations per capita participation in EU civilian crisis management.

Effectiveness and impact have been an integral part of the domestic discussion throughout Finnish participation in crisis management. Several strategic documents have stated that Finland should focus its participation on missions and operations that are the most effective from the perspective of Finnish aims in foreign and security policy (UM 2009; VNK 2014; VNK 2021). It is generally understood that evaluating and measuring effectiveness is needed, because it provides a forum for identifying lessons from past and current operations, and, therefore, helps to steer peace operations towards more impactful and effective ways of operating. Evaluations are also irreplaceable in establishing accountability of peace operations. Simply put, properly conducted evaluations lead the way to developing better, more effective operations and missions.

Considering this, it is surprising that no systematic method or practice to evaluate the effectiveness of Finnish participation in crisis management has been utilised so far, despite several governmental strategies and initiatives stating both the need as well as potential ways and methods to do this. In addition, the few evaluations that have been conducted have mostly been rather limited in scope and criticised as being vague.

This paper seeks to provide an answer to why measuring effectiveness or impact has proved to be so difficult nationally, even though its importance has been emphasised on numerous occasions. Through our analysis, we discovered that there are three interlinked issues related to measuring effectiveness that should be addressed to establish a sustainable basis for evaluations. First, the terminology and concepts used in discussions about the effectiveness of crisis management are used inconsistently. Second, the objectives set for contributing to crisis management missions and operations are either missing or imprecise. Finally, there is no clear process for, or ownership of, assessing the effectiveness of peace operations as a whole, nor the effectiveness of national contributions to these operations. Together these issues amount to systematic assessments of effectiveness not being conducted, and therefore decisions about further participation and new operations being established with limited information about past successes and failures.

We suggest three practical steps for improving the situation: creating a shared understanding of concepts, setting precise objectives, and assigning responsibility for evaluation processes. This paper is based on analysis of the use of the concept of effectiveness in Finnish discourse on crisis management, combined with a meta-analysis of recent reports on Finland's participation in crisis management activities in Afghanistan.

We also discuss whether the Finnish examples could provide answers to why measuring the effectiveness of peace operations has proved to be so challenging, and how the insight thus gained could be used to develop thinking and practice for example in the context of the EU's Common Security and Defense Policy (CSDP). We claim that there might be unrecognised or consciously disregarded aspects in the very nature of crisis management, related to its multiple level goal setting, which need to be recognised in order to fully understand the phenomena and the evaluability of its effectiveness.[1]

2 Step one: Creating a shared understanding of concepts

Regardless of the theme or topic, a shared set of terms is key to avoiding misunderstandings and misconceptions. This is especially true in the case of evaluating the effectiveness of peace operations,[2] as there seems to be many ways to understand and use key terms related to it, such as "effectiveness", "impact" and "efficiency". These concepts are interrelated, yet their relationship needs to be more clearly defined.

"Effectiveness" can be understood in multiple ways. In the context of evaluations, effectiveness refers to the extent to which an actor has reached certain set objectives in relation to the means at its disposal. However, when the concept is taken out of the context of evaluation and set in everyday language, its meaning expands toward various political and governance related definitions. A broader definition of effectiveness turns it into an umbrella term encompassing concepts such as "impact" and "efficiency", which relate to the broader topic of effectiveness or to a greater effort to discuss effectiveness.[3]

---

1 Currently there is a broad discussion on the overall need to rethink and improve international crisis management and peace operations, e.g. through applying a systems approach. For this discussion see, for example, De Coning 2018; Day 2022 and Tykkyläinen et al. 2023.

2 The evaluated "action" itself also requires some terminological decisions. In this paper we use the term "peace operations" as an umbrella term to refer generally to peacekeeping and civilian or military crisis management. However, when we refer specifically to Finnish or EU contexts we use the term (civilian or military) "crisis management", which is the term used in the EU context.

3 For example, in the names of organisations and such, e.g. Effectiveness of Peace Operations Network (EPON) or Improving the Effectiveness of Capabilities of EU Conflict Prevention (IECEU), in which cases the use of the term effectiveness is broader than the one limited to evaluation.

"Impact", on the other hand, refers to the long-term positive and negative effects or changes that result from actions and outcomes. "Efficiency", in turn, is simpler, and describes the cost-effectiveness of an action. The level of uncertainty about these concepts seems to be at its greatest in the field of peace operations. Their use is more consistent and there seems to be a much stronger shared understanding of their meaning in the domains of humanitarian assistance and development cooperation (see for example OECD 2023).

Translating evaluation and effectiveness terminology into various languages within national contexts produces additional problems. In the Finnish language, for instance, the terms "effectiveness" (vaikuttavuus) and "efficiency" (tuloksellisuus) are often used synonymously. Although these concepts are closely connected to one another in Finnish as well, we claim they should be used with more precision in the context of evaluations. The inability to do so has often led to interpretations where a theory of change is oversimplified. A typical example of this is shortening an *input–output–outcome–impact* theory of change chain into *efficiency = effectiveness.* This can lead to efficiency being interpreted as sufficient proof of effectiveness.

In Finland, the effort to make a clearer distinction between different terms has been the separation of *societal effectiveness* and *efficiency of actions*. These have been used in the Finnish governmental performance guidance manual to draw a difference between being merely efficient, and being both efficient and effective or having an impact (VM 2005, p. 25-26). The important distinction between the two is that an organisation can be efficient in its actions but still lack in the effectiveness of its results.

In some cases, the difference between impact and effectiveness has also been attributed to the type, or more specifically the author, of an evaluation. The Finnish Ministry for Foreign Affairs (MFA), for example, often refers to *impact evaluation* when the evaluation is done by external actors or consultants, and to *monitoring of effectiveness* in the case of internal assessments. These divisions between both impact/effectiveness and evaluation/monitoring are commonly used in the context of development cooperation evaluations (see for example MFA 2015).

3 Step two: Setting clear objectives

In addition to a shared understanding of key concepts, the assessment of peace operations requires standards against which to evaluate them. Pre-set goals are essential to assessing the effectiveness of peace operations. Without clear objectives and a baseline for comparison, it is very hard to evaluate whether an intervention has done what it was intended to do. This being said, interventions can, and often do have unintended consequences that should not be overlooked in evaluations. Actually, some entities, such as EPON,[4] take

4 EPON (The Effectiveness of Peace Operations Network) is an international network of over 40 partners, which aims to assess and enhance the effectiveness of international peace operations through collaborative research. For more information on EPON and its methodology see https://effectivepeaceops.net/

unpredictability, constant change and elusiveness of effectiveness as a starting point, setting evaluations of peace operations apart from humanitarian and development-focused evaluations, which tend to focus on shorter-term, project-type activities.

Different evaluation projects, even within the realm of peace operations, have approached pre-set goals in slightly different ways. The EPON methodology, for example, looks at two sets of goals: those set in the mandate of the operation, and those with a more general intention of affecting the political and security dynamics in the conflict system. Thus, the EPON evaluations have a clearly stated source for the goals they assess. The IECEU project,[5] on the other hand, looked at effectiveness by assessing it through six themes: planning, cooperation, competences, comprehensiveness, technologies and operational capabilities. By using these set categories (formulated by the evaluation team and not derived from the operations' own objectives) the IECEU project distanced the evaluation of effectiveness from the need for pre-set goals to assess, and instead looked at the effectiveness of peace operations from a non-goal-centred viewpoint. This strategy might be called for in some cases, but it should not lead to overlooking the importance of goals. After all, goals represent the reasoning for the entire peace operation, or national contribution to it.

We found a consistent lack of precise and evaluable goals for participation in crisis management when looking at the Finnish context. The goals set for participation in peace operations are often rather vague and high-level. The Finnish National Strategy for Comprehensive Crisis Management (VN 2009, 16) states that Finland's participation should focus on areas that are crucial for the goals of Finnish foreign and security policy. However, the latest government programme states that those goals are "to assume a share of the responsibility for maintaining international peace and security and to develop the capabilities and preparedness of the Defence Forces" (VN 2019).

Even in documents closer to the operational level, the objectives are often too vague to measure, such as "enhancing anti-terrorism efforts" or "improving military capabilities". These may be desirable outcomes of crisis management, but they are either very hard to assess or they can be interpreted so widely that they are virtually impossible to miss. In terms of evaluating effectiveness, these kinds of vague or broad statements make poor goals. Another example of poor setting of objectives is looking only at inputs, which are often used as 'substitutes' for goals (e.g. the desired number of national secondees in peace operations). Even though these kinds of goals are quantitative, and thus easily measurable, they should not be treated as effectiveness objectives, but rather as something that can facilitate reaching the yet-to-be-defined objective.

Another relevant issue in terms of goal setting is to acknowledge openly that peace operations and national contributions to them can be, and often are, based on goals

5 The IECEU (Improving the Effectiveness of Capabilities in EU Conflict Prevention) was a project funded by the EU's Horizon 2020 programme. It aimed to enhance the EU's conflict prevention capabilities. The project ended in 2018. (For more information on IECEU see https://ieceu-project.com.)

that are set at different levels. The operations have their goals, but nations contributing to these operations also have their own national-level goals, which may differ from the goals of the mission. In addition, both international and national goals can be more political and instrumental (e.g. providing opportunities for international military cooperation) or societal (e.g. improving the capabilities of local authorities). Thus, it should always be clearly stated whose goals are being assessed in any given evaluation. This should naturally be reflected already in goal setting, so that the national goals do not contradict or hinder the fulfilment of the operation's goals, but derive from them within the limits of the mandate. Both national as well as operational goals should be evaluated, either separately or jointly, depending on the need. However, as we argue next, neither level should be completely overlooked in evaluations.

4 Step three: Establishing a process and assigning responsibility

The third issue we discovered in the national context was a lack of a systematic evaluation process. Instead of thorough evaluations, the only regular statements about the state of Finland's contribution in relation to its goals is made in biannual crisis management reviews given to parliament. These crisis management reviews state the quantity of Finnish contributions to both military and civilian crisis management missions and operations. In the case of military operations, the reviews also include information about the operations' general goals, but the section of civilian crisis management does not even include information about all of the missions Finland is currently participating in (see for example MFA 2021). The reviews have been called out as insufficient in their capacity to provide in-depth information on effectiveness for decision-makers on multiple occasions (see for example UaVL 2018; UaVL 2020). Yet no significant improvements in their quality have been made. More comprehensive evaluations have only been conducted in a sporadic manner with no consistency in methodology or ownership, as is also evident in the reports on Afghanistan presented as an example case below. The lack of systematic evaluation processes has been addressed in previous national and international projects, such as KRIHAVA (2012) and IECEU (2018), but findings or suggestions from these have not been implemented (Ruohomäki & Hakanen 2022).

National and multilateral perspectives to assessing crisis management are obviously different. However, similar problems related to a lack of conceptual clarity and stated political aims for crisis management engagement, which we identified in the Finnish context, are also present at the EU level. When assessing effectiveness from a national perspective, part of the difficulty arises from the fact that the aims and objectives of missions are planned beyond the national context. Assessment of the missions' effectiveness at the EU level should not be a problem, as there is ample documentation stating the goals and main tasks of the missions and operations. Intervention mandates state the aims of missions, the operational and implementation plans operationalise the aims, and

the political-level goals are stated in the documentation connected to each mission. In theory, this framework gives a good basis for assessing the impact and effectiveness of the EU's Common Security and Defence Policy (CSDP) operations. In reality, however, the mandate and the political-level documents do not translate easily into measurable goals.

The Operational Plans (OPLAN) and Mission Implementation Plans (MIP), on the other hand, are designed to be measurable, but they focus on producing data for the performance assessment of the mission. More specifically, how the mission aims to fulfil its goals and at the same time use its resources. This typically leads to the production of quantifiable variables such as the number of training days that tell us something about the efficiency of the mission, but often too little about its effectiveness. The impact assessment, meaning *how* and not just *what* the mission is doing to reach its goals of stabilisation, democratisation or reform, are less clearly defined. For these, in the case of the EU, there is the process of strategic reviews. The purpose of the strategic reviews is to assess the overall progress of a given mission in order to make necessary amendments to the mandate and ultimately, to provide advice in making the overall judgement on the success and continuation of the mission. As the strategic reviews are not public documents, it is impossible to assess their quality and content, as well as how much their advice influences the political decision-making. However, it has been argued that the review processes often include pitfalls, such as polishing up results and lack of transparency (Karjalainen & Savoranta 2021). It is also questionable how productive it is to divide the evaluation of effectiveness into separate projects, with one assessing the technical performance of the mission, in terms of resource allocation etc. and the other at the political impact.

Even though EU crisis management has been operational for quite some time, no kind of shared system for evaluating the missions exists, even at the level of technical performance. The missions have different systems of assessing the efficiency of their activities, such as indicators on performance in relation to the MIP, as mentioned above. These might include "traffic light" models or other similar monitoring methods, which give the mission's leadership information on how the mission is functioning from an *internal* perspective. Thus, following and keeping track of these kinds of indicators would be better described as *monitoring internal effectiveness* than *evaluating external effectiveness.*

The methods discussed above do not assess the actual societal effectiveness, which the EU is calling impact, nor is that their intention. Rather, effectiveness is assumed to follow if the activities planned in the OPLAN and the MIP are executed as intended. The theory of change models are thus reduced to a form that is too simplistic (efficiency = effectiveness). In this respect, the assumptions are similar to those observed in the Finnish context.

This internal level of *efficiency analysis* would perhaps be the easiest to standardise as the structures and cycles of information production, such as the six-monthly-reporting period, are already in place. However, no notable work on standardisation of internal assessment has been introduced publicly. Currently there is no shared conceptual under-

standing of the concepts of evaluation, assessment, effectiveness and impact amongst the European External Action Service (EEAS) or the member states. This makes the quality of the internal assessments questionable, and definitely non-comparable. This room for improvement has been identified in the so-called civilian CSDP Compact signed by the member states in 2018, which states in its commitment no. 13 that a system of assessing operational impact needs to be designed (Council of the European Union 2018). Despite the commitment in the Compact process, such a system does not yet exist, though it is desperately needed. The new civilian-CSDP Compact, which is been drafted could provide a needed step forward.

As mentioned earlier, it is crucial that the agent or institution tasked with evaluations is clearly defined and has ownership of the process. In some cases, e.g. in the field of development cooperation, external evaluations are considered the norm and are often demanded by funders (see for example MFA 2023). This is mainly because external evaluations are more likely to be independent, less biased and more ambitious than evaluations conducted by the same institution whose actions are under scrutiny. Assessing one's own actions can, in some cases, lead to selective assessments, where effectiveness is only viewed from a narrow point of view and external effectiveness or negative effectiveness is not included in the evaluation.

However, completely externalising assessments of internal effectiveness or even societal effectiveness of the mission might not be practical for a number of reasons. First of all, the key documents needed to make these assessments are internal to the mission, as is the knowledge of these functions. Secondly, in-house evaluations may have a clearer path to contributing to decision-making and lessons-learnt processes. One possibility could therefore be to collect and analyse information internally, but in a more standardised way, and use it to conduct more robust internal evaluations and effectiveness monitoring. This information could then be redistributed to a possible agency that is at least partly external, which is responsible for conducting an independent, more political- and societal-level assessment of effectiveness. In essence, this would mean developing and standardising the current processes of internal monitoring and evaluation and creating a separate but connected and parallel track for external evaluations.

In the case of the EU, these kinds of evaluations could be executed by EEAS or member states themselves, for example by the EEAS structures, and/or by pooling expertise from the member states governmental institutions capable of conducting such analysis. Another option could be to outsource the external or societal effectiveness evaluations to an entity such as EPON. An internal evaluation would have the benefit of being able to use the restricted documents available for governmental institutions, whereas the latter, completely externalised evaluation would provide more objectivity, but would be in danger of being cut off from the decision-making processes. In both cases, these political impact assessments should follow the same standardised methodology and they should be as transparent as possible. They should also have pre-set processes to which they feed information.

5 Reports on Finnish contributions in Afghanistan as an example of ex-post evaluations

The collapse of the Western-backed government in Afghanistan in 2021 launched a worldwide quest to find answers to the seemingly total failure of international interventions. In Finland, this resulted in several reports being commissioned to evaluate Finnish participation in Afghanistan from different angles. Three reports have been published at the time of writing this paper and at least one more will be published during 2023.[6] The three released reports provide an interesting example of illustrating how a lack of precise goals and general vagueness of effectiveness terminology results in serious difficulties. The Afghanistan reports also clearly demonstrate the need for a systematic process for evaluations. The reports are unanimous in stating that there were no measurable or evaluable goals set at the national level when Finland joined the peace operations in Afghanistan. However, they do identify effectiveness and success related to the role and agency of Finland in the international system.

Reading these reports also reveals a meta-level of "analysing the analyses", which sheds light on the particular nature of crisis management as an object of evaluation and assessment. Goal setting in crisis management can be seen as a complex process, in which the most important goals might not be the ones found in mission mandates or official documents. Instead, goal setting is a work of different, changing sets of agents, which themselves are in constant change. This could be an important factor explaining why consistent and systematic assessment of crisis management has been so rarely attempted.

Three reports – three baselines

In 2022 three different reports on the Finnish participation in Afghanistan were published. Two of the reports were drafted by the government bodies, the MFA and the MoD, which focused on their domains in the government. The third one was compiled by the Finnish Institute of International Affairs (FIIA), by the commission of the Finnish Parliament and had a more overarching view.

The Afghanistan report drafted by the Finnish Ministry for Foreign Affairs (MFA 2022), delivered to the Parliament, was necessitated by the Finnish constitution, regulating the reporting on the participation to crisis management operations abroad. The report admits that effectiveness was not a priority in goal setting, while decisions were

6 The Finnish MoD and MFA as well as FIIA published their reports by the end of 2022. A fourth report commissioned by the MFA, MoD and MoI as part of the government's analysis, assessment and research activities (VNTEAS) is to be released in 2023. CMC Finland is part of the consortium working on the fourth report on the effectiveness of Finland's contribution in Afghanistan (for more information see https://tietokayttoon.fi/-/suomen-afganistan-operaatioiden-vaikuttavuus-ja-systeemisen-arviointimallin-kehittaminen).

made about Finnish participation in crisis management (MFA 2022, p. 9). It then proceeds to construct these goals by combining and interpreting different policy approaches and aspirations. This results in a four-pillar structure of higher-lever effectiveness goals related to Finnish participation: 1. Stabilisation of Afghanistan's security situation, 2. Supporting reconstruction, sustainable development and human rights, 3. Improving international and national security, and 4. Strengthening Finland and the EU as international agents. The said MFA report, does not claim, however, that the actions taken in Afghanistan or their consequences are the result of a coherent strategy.

We find this problematic for the following reasons: Being that the goals are constructed via interpretations, we cannot be sure whether these interpretations correspond to the actual processes and intentions that were present in the past. As the absence of measurable, or even identifiable, goals is recognised, it follows that evaluating strategic and operational planning is difficult. Tangible results and impact achieved might well exist, but they would not necessarily correspond with what was hoped or aimed for.

The report produced by the Finnish Ministry of Defence (MoD 2022) was the first of three Finnish reports on Afghanistan published in 2022. The absence of measurable or operationalised goals is not identified as problematic to the evaluation in the MoD report. Nonetheless, the report states that Finnish participation in Afghanistan improved Finland's ability to cooperate with its Western allies, supported Finland's reputation as a reliable partner and, therefore, concludes that it benefitted the process of Finland applying for NATO membership in 2022 (MoD 2022).

The said MoD report is a clear ex-post legitimation of Finland's military participation, which draws theory of change models that are not up to analytical scrutiny. At the same time, it reveals a difference in thinking about effectiveness and goal setting in comparison to the other two reports on Finnish contributions in Afghanistan analysed for this paper. In the case of military contributions, it seems that participation in international crisis management can be instrumental and highly effective from a national perspective, even if the specific strategic-operational goals of the given operation were left completely unfulfilled. The effectiveness produced by participating can be (and in this case is) impactful towards the capabilities of the national agent itself, recognised ex-post, without clear ex-ante goal setting.

The MoD's Afghanistan report is a representative case of an evaluation that looks at *operational efficiency* and internal developments, whereas *societal effectiveness* from a local perspective does not seem to be a priority. Nevertheless, general public and decision makers read both types of evaluations through the same lens: was it worthwhile contributing to these operations? In contrast to evaluations focusing on societal effectiveness, assessments focusing on operational efficiency steer away from a rather crucial question: At what cost to the locals were these capability improvements achieved? Commitment to ethical norms, such as the do no harm principle, also require these kinds of assessments.

The report published by the Finnish Institute of International Affairs (FIIA) (Mustasilta et al. 2022) on Finland's contributions in Afghanistan has a similar approach to goals as the MFA (2022) report. It strongly recognises that goals for national contributions to peace operations in Afghanistan were not clearly set before or even during the time that resources, both personnel as well as financial, were sent to Afghanistan. The FIIA report draws from numerous interviews and various documents to reconstruct the narratives and motivations behind Finland's participation. From these sources, the writers identify three different frames of motivations that guided Finland's contribution in Afghanistan: 1. Finland as a benefactor, 2. Finland as a partner, and 3. Finland as a learner.

While the first frame was stated in documentation and presented to the public as Finland's main reasoning for participating in the collective efforts in Afghanistan, the second frame, Finland as a partner, is presented, in hindsight, as the main driver behind decision-making. Though it was important for Finland to be seen as a responsible actor, it was even more important to use Afghanistan as an arena in which to indicate and prove allegiance to and solidarity with the United States and other international partners. The report states that in the first stages of the operations, the different motivations guided Finland to take part in some sort of harmony, but later on, these motivations contradicted each other (Mustasilta et al. 2022). This is presented in the report as displaying the lack of a comprehensive or systematic approach on a national level, as well as a lack of public recognition of the true reasoning or aims of the Finnish contribution. Unlike the first two reports, the FIIA report can be considered an external evaluation, which might partly explain its more critical viewpoint in comparison to the other two internal evaluations.

From reading these three reports describing and analysing the Finnish contributions in Afghanistan, and seeing the wide range of interpretations the reports are drawing from (which are at least partly the same material and the same events), it becomes evident what the lack of commonly shared terminology, absence of pre-set objectives and no systematic processes leads to or enables. Politically speaking, the looseness of goal setting and a limited understanding of the concept of societal effectiveness makes it possible to frame operational efficiency as effectiveness. This, in turn, allows for the favourable interpretation of the results of engagement and from the viewpoint of the present, as was the case in the MoD report. However, from the point of view of strategic planning and accountability this is clearly not sufficient. Moreover, the lack of a systematic process or clearly assigned responsibility for and ownership of evaluations has enabled a situation in which actors who have been responsible for the political guidance of the assessed activities conduct evaluations internally. The MFA and MoD reports are examples of this. Self-evaluations can be useful and called-for in view of adaptivity as well as organisational learning and development, but they should not replace external evaluations.

6 Layers of goals

The above analysis of the reports produced on the Finnish contributions in Afghanistan raise the point of instrumentality of national crisis management participation, which also affects the way we should approach the question of measuring effectiveness. The Finland as a partner narrative found in all of the reports is interesting, as the aim of partnership originates strongly from national interests towards the security of Finland itself, namely from the desire to move closer to NATO, the United States and Germany. However, this narrative is not limited to Finnish participation in military crisis management in Afghanistan. It covers the whole of Finnish participation, including civilian crisis management, which is portrayed in the reports as a part of a larger partnership strategy, but presented in public as having much more altruistic goals, such as implementing the UN's Women, Peace and Security Agenda as well as enhancing human security in Afghanistan. Nevertheless, based on the three above-mentioned reports, these goals were secondary, and can be understood as instrumental to the primary drivers of Finland's own security interests.

The existence of hidden drivers and instrumentalisation of national contribution raises acute ethical questions, which need to be addressed and acknowledged in future decisions to participate in crisis management and peace operations. Planning and assessing peace operations should also consider this. There will always be political goals and drivers, the assessment of which will be extremely difficult in practice. Goals, such as increasing the leverage Finland has in the international system, might best be judged by the democratic institutions in the form of parliamentary debates and ultimately elections, but they can also be analysed as part of an evaluation of effectiveness. Not all political goals are beyond measurement, and their existence does not mean that we can only measure and evaluate the technical aspects, or the efficiency of operations. Political goal setting could and should be as explicit and concrete as possible. This was not the case in Finnish participation in Afghanistan. In cases where assessing positive effectiveness (i.e. whether the goal has been reached) is not practically possible, the minimal requirement should be to assess the possible negative and unintended effects that occurred at the local level while pursuing political goals.

The fact that participation in peace operations carry so many secondary drivers and goals (which might eventually become primary), and which are external to the theatre of operations, may help in explaining why measuring effectiveness is so rarely put into use in peace operations. Political drivers and goal setting are integral to the nation state, or the multilateral organisation of nation states, but not so in a world of assessing humanitarian and development work, with its more practical and shorter project-type time frames. Would this difference explain why evaluations have been widely used only in the latter? The EU is an interesting example of this, where its Commission-run project-type activities are regularly evaluated, while the member state-driven CSDP activities are not. Could this be because of the multiple levels of overt and covert goals set by the member states and the EU as a whole respectively?

As the decision-makers in peace operations are aware of the unwritten drivers and goals, sometimes even constituting participation, they have been slow to commission assessments that would either miss these goals or reveal their existence in a less favourable light. The national example of Finland would support this reading. Despite numerous proclamations, projects and discussions spanning over a decade, no systematic structures of assessment have been put in place. On the other hand, assessments are being commissioned and there seems to be a legitimate desire to validate participation in peace operations. However, instead of primary drivers, effectiveness is viewed through secondary, often less political actions and goals. Arguably, there is a desire to see whether the (covert) political aims of participation have been met, even though the assessments could not be directed towards those goals.

This tendency to overlook the primary drivers becomes obvious when looking at both evaluation commissions as well as projects aimed at developing evaluation methods or processes in the national context. There is often only a very limited need to justify the crisis management operation itself at the national level. Instead, what is sought after in the evaluation commissions, is a *justification for the national contribution*. Therefore, while the commissions often state a will and a need to examine the *effectiveness* of *participation* in crisis management in light of the mandate of the operation, they simultaneously seek an evaluation that isolates the effect of the national contribution from its international setting. This is often not possible or even meaningful because national contributions to a multinational crisis management operation are, more often than not, just small parts in different corners of a much bigger puzzle. Thus, many of the assigned evaluators turn to assessing national goals for participation instead of the strategic and operational goals of the whole operation.

In order for the evaluation processes to provide the information requested in the commissions, a better understanding of this is needed on the side of the commissioners, which are often ministries. Simply put, one may evaluate the overall effectiveness of an operation by looking at the international input, or evaluate the effectiveness of national participation in the light of national goals. The national goals may, and often to some degree do, overlap with the mission's mandate, but they are rarely identical. What follows is that separate evaluations are needed on both the national and international level. In both cases shared concepts, better goal setting and clear responsibility for and ownership of an evaluation process are needed.

7 The way forward

What follows from all of the above is that if we want to improve evaluations, **we first need a shared understanding of key concepts**. In theory, this is rather simple, as it is a matter of explaining the meaning behind the used terminology instead of just assuming that everyone is on the same page. In practice, however, this might prove to be slightly more challenging. While measuring operational efficiency and effectiveness in relation to internal developments is often valid and even called for, it should not be framed as an evaluation focusing on societal effectiveness or impact. The difference between the two should be clearly stated. It is also vital that **when evaluating effectiveness in relation to national or political goals, the effects on the local level should not be completely disregarded**. Evaluations are needed on both levels, and therefore responsibility for the evaluation of societal effectiveness needs to be clearly assigned.

In an ideal situation, every peace operation would have clear, realistic and assessable operational goals for how it is going to defuse a conflict situation, or how it is going to help civil society. The goal setting for operations would ideally be based on a solid conflict analysis and substantial consultations between various local stakeholders, such as government, the general population and minority groups. Additionally, every country contributing to the operation would have its own similar processes to make sure that participation is assessed as objectively as possible. However, the nature of peace operations is such that this is not possible in many cases. In the real world, one has to seek a balance between shortcomings in terms of information, consultation, goal setting, planning and evaluation. Nevertheless, measurable goals, at least to some extent, need to be found as soon as possible after the operation has been launched, if not before. In the former case, it might be perfectly justifiable to launch, learn and then adopt realities in the field and adjust the goals accordingly. In any case, goals are needed, even if they cannot be formulated from day one.

The political goals of peace operations are much more difficult to articulate and submit for assessment than operational ones. Political goals can be discreet from the perspective of the multilateral organisation, the host country or an individual country seconding personnel for the operation. But these goals do exist, as we found when reading the Finnish analysis of the participation in crisis management activities in Afghanistan. It is up for debate whether these types of goals could be subjected to assessment. Clearly some of them can be assessed, but some of the political goals, especially if they are particular goals of a seconding nation state, might be best left for the political processes of that country itself.

However, the politico-strategic goals that the whole operation shares need to be openly discussed and assessed as part of normal democratic oversight. **What is most important to understand, is that the goal setting for a peace operation will include goals at various levels.** Some of them can be easily quantifiable, but others, especially political goals, need to be assessed in a more holistic way by cross-checking whether the theory of change used and the tools utilised provide the best possible result. A clear-cut assessment of whether the mission has been a success or not is probably not possible; instead, the question is about weighing the impact with the resources used, while not forgetting ethical principles such as doing no harm and the need for conflict sensitivity.

Whichever way the above-mentioned questions are solved, a successful evaluation will need structures. **The act of assessment itself must be mandated to a specific body**, whether it lies within the structures of the government or a multilateral agent (while separated from the chain of command) or whether it is completely external to the action. In the latter case, the obvious benefit is objectivity, while the possible loss is being cut off from planning and execution, which risks leaving assessment estranged from its object. In ideal situations, some kind of hybrid form might be used where the object of assessment is not completely passive, but takes an active role in continuous assessment and improvement. **Regardless of the mandated evaluating body, both the evaluation process and its findings need to be tied to the relevant decision-making processes.** The result would be an adaptive and agile peace operation, which has the possibility to self-correct within the boundaries of its mandate and that is externally assessed periodically, allowing for more fundamental changes when needed. The EU's crisis management activities, which currently lack any kind of standardised structures for evaluation, could benefit from this and create a state-of-the-art working system to start from.

References

Council of the European Union (2018). Conclusions of the Council and of the Representatives of the Governments of the Member States, meeting within the Council, on the establishment of a Civilian CSDP Compact. https://data.consilium.europa.eu/doc/document/ST-14305-2018-INIT/en/pdf

Day, A. (2022). States of Disorder. Ecosystems of Governance. Complexity Theory Applied to UN Statebuilding in the DRC and South Sudan. Oxford University Press.

de Coning, C. (2018). Adaptive Peacebuilding. International Affairs. 94:2. https://doi.org/10.1093/ia/iix251

EPON (2023). The Effectiveness of Peace Operations Network. https://effectivepeaceops.net/

IECEU (2018). Improving the Effectiveness of Capabilities in EU conflict prevention. https://ieceu-project.com/#focus

Karjalainen, T. & Savoranta, V. (2021). The EU's Strategic Approach to Crisis Management: Building a Tenet from Praxis. FIIA Analysis 2021/11.

KRIHAVA (2012). Kokonaisvaltaisen kriisinhallintatoiminnan vaikuttavuuden arvioinnin kehittäminen (KRIHAVA), Loppuraportti.

MFA (2015). Development evaluation norm of the Ministry for Foreign Affairs (MFA) (Translation from Finnish, unofficial). https://um.fi/documents/35732/48132/development_evaluation_norm_2015

MFA (2021). Suomen osallistuminen kansainväliseen kriisinhallintaan. https://www.eduskunta.fi/FI/vaski/Liiteasiakirja/Documents/EDK-2022-AK-1407.pdf

MFA (2022). Suomi osana kansainvälisen yhteisön kriisinhallinta- ja kehitysyhteistyötoimintaa Afganistanissa 2001–2021. https://um.fi/documents/35732/0/UH+AF+selvitys+FINAL+221222.pdf/3e6065a1-b072-cc9e-e369-f32a1b889450?t=1671717356679

MFA (2023). Development evaluation. https://um.fi/evaluation-of-development-cooperation

MoD (2022). Puolustusministeriön Afganistan selvitys. http://urn.fi/URN:ISBN:978-951-663-115-1

Mustasilta, K., Karjalainen, T., Stewart, T. & Salo, M. (2022). Suomi Afganistanissa 2001–2021: Vakauttamisesta ulko- ja turvallisuuspoliittisten suhteiden vaalimiseen. https://www.fiia.fi/wp-content/uploads/2022/12/fiia-report-72_suomi-afganistanissa-2001-2021_katariina-mustasilta-et-al..pdf

OECD (2023). Evaluation Criteria. https://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm

Ruohomäki, J. & Hakanen, J. (2022). Käsitteistä käytäntöön – Kohti vaikuttavampaa kriisinhallinta. Valtioneuvoston selvitys- ja tutkimustoiminnan artikkelisarja 31/2022. https://tietokayttoon.fi/julkaisu?pubid=42501

Tykkyläinen, S.,Karjalainen, A., Brusset E., Hario, P. (2023) Systems Approach to Peace Operations. https://www.cmcfinland.fi/wp-content/uploads/2023/02/FINAL_CMC-Systems-approaches-to-peace-operations-WEB-versio.pdf

UaVL (2018). Valtioneuvoston selvitys Suomen osallistumisesta YK:n UNIFIL-operaatioon Libanonissa. https://www.eduskunta.fi/FI/vaski/Lausunto/Sivut/UaVL_1+2018.aspx

UaVL (2020). Valtioneuvoston selvitys: Suomen osallistuminen kansainväliseen kriisinhallintaan. https://www.eduskunta.fi/FI/vaski/Lausunto/Sivut/UaVL_1+2020.aspx

UM (2009) Suomen kokonaisvaltainen kriisinhallintastrategia, s. 6. Ulkoministeriö, Helsinki. https://um.fi/documents/35732/48132/suomen_kokonaisvaltainen_kriisinhallintastrategia__pdf_

Valtioneuvosto (2021) Vaikuttavaa kriisinhallintaa: Kriisinhallinnan parlamentaarisen komitean suosituksia Suomen kriisinhallinnan kehittämiseksi. Valtioneuvoston julkaisuja 2021:13 http://urn.fi/URN:ISBN:978-952-383-848-2

VM (2005). Tulosohjauksen käsikirja. https://vm.fi/documents/10623/307545/Tulosohjauksen+k%C3%A4sikirja.pdf/b7f9a7f9-2b46-4dbb-bb66-85bf1074b88a

VN (2019). Inclusive and competent Finland – a socially, economically and ecologically sustainable society – Programme of Prime Minister Sanna Marin's Government 2019. https://valtioneuvosto.fi/en/marin/government-programme

VNK (2014) Siviilikriisinhallinnan kansallinen strategia. Valtioneuvoston kanslian julkaisusarja 8/2014. http://urn.fi/URN:ISBN:978-952-287-123-7