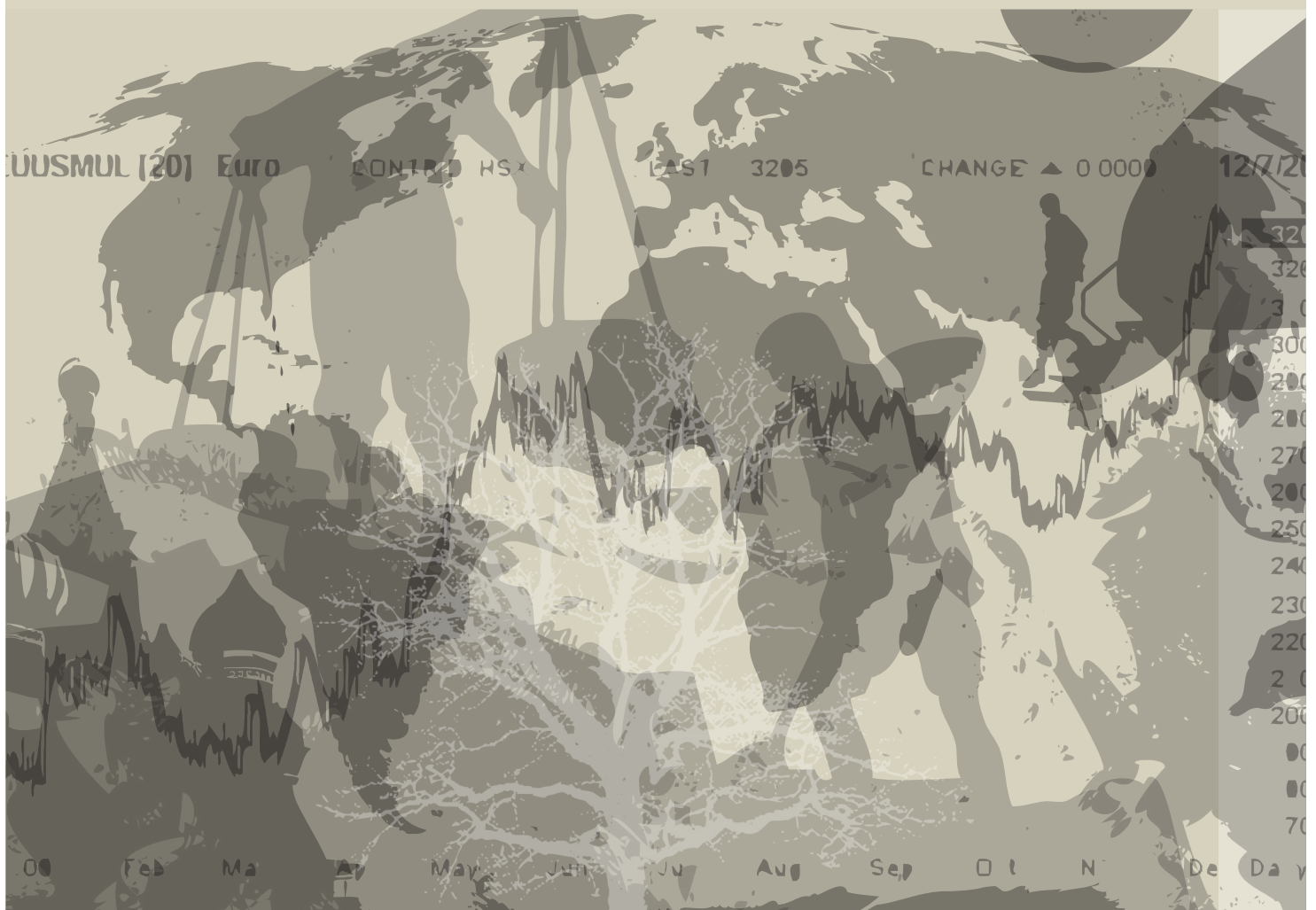


Measuring the Success of Peace Operations: Directions in Academic Literature

Mateja Peter



NUPI Working Paper 862

Publisher: Norwegian Institute of International Affairs
Copyright: © Norwegian Institute of International Affairs 2016

Any views expressed in this publication are those of the authors. They should not be interpreted as reflecting the views of the Norwegian Institute of International Affairs. The text may not be printed in part or in full without the permission of the authors.

Visiting address: C.J. Hambros plass 2d
Address: P.O. Box 8159 Dep.
NO-0033 Oslo, Norway
Internet: www.nupi.no
E-mail: post@nupi.no
Fax: [+ 47] 22 99 40 50
Tel: [+ 47] 22 99 40 00

Measuring the Success of Peace Operations: Directions in Academic Literature

Mateja Peter

Published by Norwegian Institute of International Affairs

Some peace operations are seen as ‘successful’ and others as ‘unsuccessful’, often with little discussion of the yardsticks used for measuring such success or failure. While some operations are relatively easy to put into boxes labelled ‘success’ and ‘failure’, most are not. Understanding what constitutes success and failure in peace operations is necessary for building rigorous scholarly knowledge about international interventions and their effects, and for making sound policy decisions. Despite the growth in the study of peace operations, on UN peacekeeping in particular, one important aspect has been under-researched: *how* to judge whether a peace operation has been successful or not. Evaluation studies tend to focus on the sources of success, at the expense of the criteria employed in assessing success. Researchers generally draw on their personal predispositions and ideas as to what success means, and then focus on the various factors that contribute to it. However, analytically speaking, determining what is meant by success must be the first step.

This background paper examines how the academic literature has approached the question of success in peace operations. Here it should be noted that many theoretical and methodological issues have not been settled, nor does this contribution seek to resolve them. The aim here is to shed light on the issue, and indicate where choices need to be made for research into success to be rigorous. The first section examines differences between the way that scholarly and practitioner analyses approach this question. The section that follows looks at how the academic literature has approached the definition of success and where some of the fault-lines lie. In the conclusion I outline a number of methodological decisions that need to be made when conducting research on success.

How do scholarly and practitioner analyses of success in peace operations differ?

Since the mid-1990s, when academic research on peace operations became more systematic, there have been calls for uniting theory and practice.¹ Many scholars are concerned with the possible contributions their work can make to policy. There has been a growth in research that encourages ‘a more collaborative approach in which agreed definitions emerge that serve common goals of more-peaceful societies’.² However, despite an often congruent comprehension of overall goals, scholarly and practitioner analyses of success in peace operations tend to differ on two counts.

¹ See for example, A.B. Fetherston (1994) *Towards a Theory on United Nations Peacekeeping*. New York: Palgrave.

² Daniel Druckman and Paul Stern (1999) Perspectives on Evaluating Peacekeeping Missions. *The International Journal of Peace Studies* 4(1), available at http://www.gmu.edu/programs/icar/ijps/vol4_1/druckman.htm.

First, academic research tends to focus on generalizable criteria, whereas practitioners favour criteria unique to specific missions. This is not surprising. Many scholars look for comparative aspects in order to contribute to their primary academic interest: theory building. By contrast, policy-makers take their starting point in practice and trying to improve missions; their concern is with what works and what does not in a specific operation and a specific context. That being said, with the proliferation of evaluation departments in many international organizations, also practitioners have started moving towards comparative analyses and generalizable aspects. Here they have focused particularly on factors/sources of success, examining success along one chosen dimension (for example, how successful missions are in promoting security sector reform). In this way, practitioners attempt to avoid some of the methodological issues around generalization that have dominated much of the scholarly debate.

The second distinction between scholarly analyses and those of practitioners is the difference in focus. Scholars usually favour a macro-approach to studying success, focusing on the broader political goals and longer-term consequences. Practitioners, by contrast, tend to focus on mandate implementation, whether in whole or in part. Such assessments are used as lessons learned, for fine-tuning and adjusting the operation – usually with a fairly short-term focus. Policy-makers rarely question the overall idea of peace operations or the theory of change entailed. It is assumed that achieving short-term goals is a precondition for longer-term success. For example, the success categorization developed by NATO differentiates between success on the strategic and on the operational/tactical levels, and links the fulfilment of strategic goals to successful implementation of operational/tactical ones.³ Practitioners concentrate on factors under their control, subsuming these under operational goals. Scholars, however, focus primarily on broader benchmarks of success, such as resolution of a conflict or reduction of human suffering. For that reason, they are more inclined to include in their analyses factors beyond the control of the peace operation in question.

How does the academic literature define success? What are some challenges?

Scholars have generally approached the question of success from one of two vantage points: case studies, or large datasets on conflicts and interventions. Many recent analyses combine the two.⁴ Both vantage points have their advantages and disadvantages. Case studies generate

³ NATO Allied Joint Publication, 3.4.1. *Peace Support Operations*, July 2001, available at <https://info.publicintelligence.net/NATO-PeaceSupport.pdf>.

⁴ Lise Morjé Howard (2008) *UN Peacekeeping in Civil Wars*, Cambridge: Cambridge University Press; Virginia Page Fortna (2008) *Does Peacekeeping Work? Shaping Belligerents' Choices after Civil War*, NJ: Princeton University Press.

specific in-depth knowledge about particular contexts. They do not smooth out the differences between operations. However, they often have little to say about the ‘substance’ of success that could be translated across cases – something that is needed for a robust definition. What complicates such analyses of success is the broad empirical bias towards both the unsuccessful elements of missions and operations deemed unsuccessful in their entirety.⁵

In the academic literature, most discussion about the definition of success has therefore occurred in connection with quantitative research. Unlike individual case studies, such work has attempted to find out whether peace operations as *an instrument* of third-party intervention are successful. These studies usually begin by providing a yardstick for success. However, most quantitative studies define success as ‘sustainable peace’, ‘durable peace’ or ‘absence of war’ – which is of less operational utility for diverse case studies.

Among scholars there is no agreed definition of success. As noted by Thomas G. Weiss, what makes it difficult to gauge success is ‘the ambiguity of success and failure, as well as the time frame used to measure the durability of results’.⁶ Paul F. Diehl and Daniel Druckman are fairly detailed in their enumeration of the challenges involved in evaluating peace operations, seeing these as including (but not limited to) ‘the distinction between performance or process and outcomes, multiple goals and stakeholders, developing effectiveness metrics, the distinction between problem-solving, legal, economic, and political approaches to the meaning of effectiveness, and differences between short and long-term evaluations’.⁷

Despite these challenges, many scholars have attempted to provide definitions and criteria for the success of peace operations. These have ranged from minimalist standards, focusing on mandate implementation, to maximalist ones. The most influential discussions on peacekeeping success in scholarly debates have tended to employ some version of maximalist standards, looking into the macro-question of whether international interventions improve the chances of peace.⁸

⁵ Lise Morjé Howard (2008), 2.

⁶ Thomas G. Weiss (1995) *The United Nations and Civil Wars*, in Thomas G. Weiss (ed.), *The United Nations and Civil Wars*, Boulder, CO: Lynne Rienner, 197.

⁷ Paul F. Diehl and Daniel Druckman (2015) *Evaluating Peace Operations*, in Joachim A. Koops, Thierry Tardy, Norrie MacQueen and Paul D. Williams (eds), *The Oxford Handbook of UN Peacekeeping Operations*, Oxford: Oxford University Press.

⁸ See for example Caroline Hartzell, Matthew Hoddie and Donald Rotschild (2001) *Stabilizing the Peace After Civil War: An Investigation of Some Key Variables*, *International Organization* 55(1), 183–208; Michael W. Doyle and Nicholas Sambanis (2006) *Making War and Building Peace*, Princeton, NJ: Princeton University Press; Jair van der Lijn (2006) *Walking the Tightrope: Do UN Peacekeeping Operations Actually Contribute to Durable Peace*. Amsterdam: Rozenberg Publishers; Virginia

For example, Virginia Page Fortna judges the success of UN peacekeeping in terms of whether the presence of peace operations reduces the risk of another war. Asking whether peace lasts longer when peacekeepers deploy than when they are absent, she examines the causal mechanisms through which their presence may make peace more stable – durable peace as a benchmark.⁹ Similarly, Kari M. Osland uses positive change towards a defined end-goal as a way of measuring success.¹⁰ To bring the idea of success closer to a view of the policy community, Michael Doyle and Nicholas Sambanis argue that the best standard is what they call ‘participatory peace’: a situation that involves an end to the war, the absence of significant residual violence, undivided sovereignty, and a minimum level of political openness.¹¹ While different scholars may express a preference for one or the other, there are no easy answers to how minimalist or comprehensive measurements of success should be.¹² A particularly contentious issue is to what extent the external environment outside the direct influence of peace operations should be incorporated in measuring success.

Among the earliest and most often cited attempts at providing comprehensive and generalizable criteria are Paul Diehl’s two measures of success – limitation of armed conflict, and resolution of an underlying conflict – which served as the basis for a set of indicators.¹³ This second criterion has been criticized in subsequent analyses, including case studies, because the nature of an underlying conflict may well change over time, especially since the factors that ignite wars are not the same as those that serve to perpetuate the conflict. In addition, Diehl developed his measures of success for a test of six cases of UN peacekeeping during the Cold War, raising questions over whether such criteria are applicable and useful to assess the multidimensional peace operations that have emerged since then.

Page Fortna (2008); Nicholas Sambanis (2008) Short-Term and Long-Term Effects of United Nations Peace Operations, *World Bank Economic Review* 22, 9–32; Anke Hoeffler (2014) Can International Interventions Secure the Peace? *Area Studies Review* 17(1), 75–94.

⁹ Virginia Page Fortna (2008), Chapter 5.

¹⁰ Kari Margrethe Osland (2014) *Much Ado About Nothing? The Impact of International Assistance to Police Reform in Afghanistan, Bosnia and Herzegovina, Kosovo, Serbia and South Sudan. A Comparative Case Study and Developing a Model for Evaluating Democratic Policing*. Oslo: Faculty of Social Sciences, University of Oslo/Akademika Publishing.

¹¹ Michael W. Doyle and Nicholas Sambanis (2006) *Making War and Building Peace*. Princeton, NJ: Princeton University Press.

¹² Fen Osler Hampson (1996) *Nurturing Peace: Why Peace Settlements Succeed or Fail*. Washington, DC: USIP Press, 9.

¹³ Paul F. Diehl (1994) *International Peacekeeping*. Baltimore, MD: John Hopkins University Press.

In his review of Diehl's book, Robert C. Johansen argued that success should not be measured against 'an ideal state of peace (for example, no armed conflict after deployment) or against an ideal form of conflict resolution (for example, settlement of long-standing animosities)'. He considered such measures to be 'normatively unfair and scientifically unproductive'. To evaluate the utility of peacekeeping, scholars should instead '(1) assess the effect of peacekeeping forces on local people affected by their work, and (2) compare the degree of misunderstanding, tension, or violence in the presence of UN peacekeepers to the estimated results of balance-of-power activity without peacekeeping.'¹⁴ However, such counterfactual analyses are not only difficult to conduct – they also give rise to a new set of methodological issues, a point to which I return in the concluding section of this paper.

Another early attempt at providing generalizable criteria for success that more clearly incorporated differences between missions was developed by Marjorie Ann Brown. She posits three criteria: Was the mandate fulfilled, as specified by the appropriate Security Council resolution? Did the operation lead to resolution of the underlying disputes of the conflict? and Did the presence of the operation contribute to the maintenance of international peace and security by reducing or eliminating conflict in the area of the operation?¹⁵ If Diehl was more systematic and detailed in developing the indicators for measuring the last two, Brown was clearer on the importance of different mandates.

Duane Bratt¹⁶ sought to adapt these evaluation criteria further to the realities of post-Cold War peace operations. He used three of the criteria employed by Diehl and Brown – mandate performance, facilitating conflict resolution, and containing the conflict – to assess 39 UN peacekeeping operations. Importantly, he altered and adapted some of the indicators developed by Diehl. For example, the indicator 'combatant deaths' was replaced by a more encompassing assessment that also included civilian deaths, as well as indirect 'natural' deaths that could have been prevented had humanitarian aid shipments not been blocked because of the fighting. Given the prevalence of civilian casualties in intra-state conflicts, such modifications are necessary to reflect the evolving *raison d'être* of peace operations. Moreover, Bratt's approach to the modification of earlier indicators offers lessons for how to utilize 'older criteria' and adapt them to new types of missions.

¹⁴ Robert C. Johansen (1994) U.N. Peacekeeping: How Should We Measure Success? Review of *International Peacekeeping* by Paul F. Diehl, *Mershon International Studies Review* 38(1), 309–310.

¹⁵ Marjorie Ann Brown (1993) *United Nations Peacekeeping: Historical Overview and Current Issues*. Report for Congress, Washington DC: Congressional Research Service.

¹⁶ Duane Bratt (1996) Assessing the Success of UN Peacekeeping Operations, *International Peacekeeping* 3(4), 64–81.

In contrast to Diehl, both Brown and Bratt included *mandates* in their criteria of success. They deemed such inclusion important, in view of the increasing variety of peace operations in the post-Cold War era. However, there are at least two challenges involved in the singular use of mandate performance. First, mandates are often vague, and ‘there is much room for debate on the scope and detail of the operation's missions’.¹⁷ Objectives encompassed in them, such as rule of law, reconciliation or good governance, are subjective notions with competing definitions. Second, the mandates given to peace operations are often unrealistic. Operations are given a laundry list of tasks and objectives to achieve, stemming from political considerations in the mandating body, for example the UN Security Council. There are often no reasonable expectations within the mandating body or the operation leadership that the mission would ever achieve these objectives. It is thus practically unfair and theoretically unproductive to judge them solely according to these criteria. When looking at mandates as a benchmark, subjective decisions on what to include and what to exclude need to be made.

While these early discussions on criteria remain a kind of a gold standard in scholarly discussions of success, new, more expansive, criteria developed as peace operations started including peacebuilding and statebuilding elements in their mandates. Roland Paris calls for a very high standard of success by arguing for a criterion of ‘stable and lasting peace’ supported by market democracy.¹⁸ Similarly, Elizabeth Cousens, while warning about holding international efforts to an impossibly high standard, develops three goals and benchmarks for success: self-enforcing ceasefire; self-enforcing peace; and democracy, justice and equity.¹⁹ Stephen Stedman and George Downs have proposed slightly more minimalist measures for operationalizing mission success, involving scores on two variables: whether large-scale violence is brought to an end while the implementers are present; and whether the war is terminated on a self-enforcing basis so that the implementers can go home without fear of the war breaking out again.²⁰

There is also another important distinction between the approach of Diehl and Bratt and these more recent definitions of success. Diehl and Bratt rely on a set of indicators which need to be combined on a scale in order to assess the degree of success. While the criteria are increasingly more refined, how to combine them depends on a subjective decision.

¹⁷ Paul F. Diehl (2008) *Peace Operations*. Cambridge: Polity, 123.

¹⁸ Roland Paris (2004) *At War's End*. Cambridge: Cambridge University Press, 59.

¹⁹ Elizabeth M. Cousens (2001) Introduction, in Elizabeth M. Cousens and Chetan Kumar with Karin Wermester (eds) *Peacebuilding as Politics: Cultivating Peace in Fragile Societies*. Boulder, CO: Lynne Rienner, 11.

²⁰ Stephen Stedman and George Downs (2001) Evaluation Issues in Peace Implementation, in Elizabeth M. Cousens and Chetan Kumar with Karin Wermester (eds) *Peacebuilding as Politics*, 50.

Are all indicators equally important, or should we be weighing them with regard to importance? As individual researchers might decide differently on these issues, measurements of success can differ. Stedman and Downs thus argue for a more limited number of general measures.²¹ However, these can entail their own problems of subjective assessment.

Some analyses have tried to resolve the complex challenge of having to choose between quantifying measures or qualitatively describing the level of success by combining the two. This can be done one of two ways. The first can be seen in how Darya Pushkina approaches her four criteria of success. While all of them have sub-questions (indicators), she provides an overall assessment only for each criterion, not for the indicators. Her criteria for success of peace operations are as follows: limiting violent conflict in the host state; reduction of human suffering; prevention of violent conflict beyond the object state's borders; and promotion of conflict resolution.²² In contrast, Lise Morjé Howard's two criteria: mandate implementation; and a broader assessment of the state of the country after completion of the UN intervention, are treated differently. She breaks down the first criterion along the various tasks that were assigned to the mission (indicators), while the second provides for a qualitative overall assessment.²³ Both of these studies indicate that a combination of a macro- and micro-level analysis is the preferred approach of scholars at the moment. This way analysts capture both the variation of missions and the overall goals of peace operations at the same time.

Choices that need to be made

As seen in the previous section, scholars have not come up with one definition or one set of indicators for success of peace operations. Much contention remains. Any researcher embarking on an evaluation study must make several analytical decisions as to how to operationalize success. This section summarizes four such decisions as outlined by Diehl and Druckman and offers a brief discussion of how the academic literature has approached these choices.²⁴

Stakeholders in peace operations

In defining success, we must always ask 'success for whom?' What matters in the operationalization of success will depend on the addressees, and on the purpose of the analysis. Different actors have different goals or assign different priorities to shared goals. Such differen-

²¹ *Ibid.*, 48.

²² Darya Pushkina (2006) A Recipe for Success? Ingredients of a Successful Peacekeeping Mission, *International Peacekeeping* 13(2), 133–149.

²³ See Lise Morjé Howard (2008), 7–8.

²⁴ Paul F. Diehl and Daniel Druckman (2015).

tiation is evident also within the mission, where the military and the civilian part may differ in their priorities. To a large extent, these goals can determine the choice of indicators of success. The more specific the indicators, the more clearly can we see for whom the study is conducted. Stakeholders might share some interests (for example, limiting violence), but these interests will not always be entirely coterminous. Troop-contributing countries or the mission as a whole might measure their success in terms of the goal of limiting casualties to their own personnel. However, achieving that version of success could lead these troops and mission as a whole to undermine their task of protecting civilians. As Diehl and Druckman point out, often left out of calculations are the interests of the weak: the local population in the area of conflict. It is for this reason that most scholars focus their criteria for success on macro-level concerns. The benchmarks used in scholarly analyses tend to promote the goals of the abstract international community (such as stability, limitation and resolution of armed conflict) and humanity (like human security, reduction of human suffering, justice).

Time perspective

When evaluating peace operations, the researcher needs to decide at what point to make an assessment and what time period to assess. While short-term and long-term success may be connected, meeting short-term goals does not necessarily lead to a positive impact over a longer period. As noted above, practitioners tend to focus on short-term perspectives, such as the achievement of objectives during the course of the operation or immediately after mission withdrawal. Such assessments are needed for modifying policies, adapting strategies and planning exits. However, several problems arise with short-term assessments: such evaluations can become quickly outdated; short-term trends are often misleading; and, perhaps most importantly, the very presence of the operation can distort local relations to the degree where it is impossible to assess whether its interventions are self-sustainable.

Employing a longer-term perspective can ameliorate some of these problems. Most academic scholarship has therefore adopted a slightly longer-term perspective, looking at the situation in the country a few years after the operation has left (or after it started), with the most sophisticated analyses adopting duration models. Such analyses do not have a cut-off time, but measure the length of time that the peace lasts.²⁵ However, such studies also have their drawbacks. The most difficult methodological problem is to what extent the intervening forces can be excluded from such studies. The longer the period, the more likely will it be that external factors – such as regime change, natural disaster or global economic downturn – could have a larger effect on local conditions than the legacy of a peace operation. Moreover, such

²⁵ See for example Virginia Page Fortna (2008).

longer-term evaluations are not immediately useful for policy and planning of ongoing and intended operations.

Baseline for assessment

Against what should the success of an operation success be judged? That is another dilemma that needs to be settled before commencing a research study. As noted above, judging peace operation success against some kind of ideal state of peace is unproductive. Some researchers therefore propose conducting counterfactual analysis, where achievements and failures of a peace operation are compared against a hypothetical situation where no action is taken by the international community. However, such analyses are not only difficult: they also create an extremely low threshold for peace operations to be labelled 'successful'. As noted by Diehl and Druckman, even a marginal improvement of the situation could be defined as success. More importantly, such analyses ignore the fact that the international community rarely chooses between a peace operation and inaction. The repertoire of alternative instruments available will range from sanctions and enforcement action to diplomatic endeavours and good offices. There might also be several different actors able to deploy a peace operation. Scenario-based analyses that include probable alternatives are difficult to conduct, but offer a good way of assessing the success of peace operations. Another possible baseline is to compare the 'before' and 'after' state of affairs. Here, however, we must bear in mind that the outcome of such an analysis depends greatly on whether an operation is deployed during a conflict or after a ceasefire. What the operation is compared to becomes crucial.

Some researchers have abandoned the search for an absolute measure of success altogether, opting instead to assess relative success across peace operations. Such analyses do not seek to produce absolute measures of success and thus circumvent some of the methodological issues indicated above. They do, however, highlight the importance of the last decision: that on what to include in the 'population of peace operations'.

Types of peace operations

Comparative analyses are best conducted between like and like. With an increasing heterogeneity of peace operations, finding such missions has become increasingly difficult. To differentiate between missions, most recent analyses include mission mandate as a guideline in determining success. However, blindly following the mandate as a benchmark for success is highly problematic. The analyst should step back and identify the primary *goals* of an operation. While mandates might differ between missions, the primary goals are more general. For example, in her analysis Page Fortna groups missions under consent-based and peace enforcement categories.²⁶ On the other hand, Sambanis dif-

²⁶ *Ibid.*

ferentiates between ‘facilitative missions, which provide monitoring and reporting (observer and traditional peacekeeping operations), and transformational UN missions (multidimensional, enforcement, and transitional administration)’,²⁷

Such goals also tend to coincide with the goals of the wider international community and humanity. When scholars have chosen to focus exclusively on primary goals they have judged success on these macro-level concerns. Others, however, have also looked at the more specific objectives of the mandate and based on them provided multiple questions/indicators. The first approach, with the focus on primary goals and macro-level concerns, enables greater comparability, whereas the second approach provides for more detailed, mission-specific analyses.

²⁷ Nicholas Sambanis (2008), 14.



Norwegian Institute of International Affairs

Established in 1959, the Norwegian Institute of International Affairs [NUPI] is a leading independent research institute on international politics and areas of relevance to Norwegian foreign policy. Formally under the Ministry of Education and Research, NUPI nevertheless operates as an independent, non-political instance in all its professional activities. Research undertaken at NUPI ranges from short-term applied research to more long-term basic research.

About the Author

Mateja Peter is Lecturer at University of St Andrews and Senior Research Fellow at the Norwegian Institute of International Affairs (NUPI). This paper has been prepared with the support of the Training for Peace programme at NUPI.

NUPI

Norwegian Institute of International Affairs
C.J. Hambros plass 2D
PO Box 8159 Dep. NO-0033 Oslo, Norway
www.nupi.no | info@nupi.no