The Performance of International Organizations:

A Policy Output Approach

Jonas Tallberg, Thomas Sommerer, Theresa Squatrito, and Magnus Lundgren

Abstract: Many problems confronting today's societies are transnational in character, leading states to increasingly rely on international organizations (IOs) for policy solutions. Yet the performance of IOs varies extensively. This article suggests that systematic, comparative research is required to advance our understanding of IO performance, and that a policy output approach offers particular advantages for that purpose. This approach privileges the results IOs produce in terms of policies, and is distinct from the main alternative approaches to IO performance, emphasizing either behavioral change by targets (policy outcome) or problem-solving effectiveness (policy impact). The article introduces a typology that captures five generic features of policy – volume, orientation, type, instrument, and target – and explains how these dimensions may be used to map the output and assess the performance of IOs. The article concludes by discussing what methods and theories may be most useful in explaining policy output, as the research agenda on IO performance advances.

Introduction

Many problems confronting today's societies are transnational in character, leading states to increasingly rely on international organizations (IOs) for policy solutions. Yet the performance of IOs varies extensively. While some IOs are highly successful in developing and enforcing policy, others are less successful. Whereas the European Court of Human Rights is famous for its strong record, for example, the United Nations (UN) Human Rights Commission was tainted by poor performance. While the Council of Ministers of the European Union (EU) sustains a high decision-making speed even after Eastern enlargement, the UN Security Council is notorious for its inability to reach agreement. How can we account for this mixed record in IO performance?

Scholars have recently begun to explore the sources of IO performance (Gutner and Thompson 2010). Yet most existing work consists of case studies of individual IOs (e.g., Elsig 2010; Lipson 2010; Pollack and Hafner-Burton 2010; Weaver 2010; Oberthür et al. 2013). While informative for the particular case, these studies tell us little about general patterns and explanations of IO performance. To advance our understanding of IO performance, we need more systematic analysis across a large number of cases. This requires a conceptualization of performance amenable to comparative, large-N analysis.

The purpose of this article is to offer such a conceptualization. We advance an output-based approach to IO performance, privileging the results IOs produce in terms of policies (cf. Easton 1965). Policy output is a requirement for IOs to change actor behavior and reduce societal problems. Also, policy output is more readily within the control of IOs, and easier to measure and compare across organizations, than alternative conceptualizations of performance.

Specifically, we propose that IO performance is best conceptualized as the adoption of policy output that matches the underlying cooperation problem(s) that an organization is mandated to address. In developing this conceptualization, we introduce a taxonomy of policy output with five generic dimensions: volume, orientation, type, instrument, and target. All dimensions are relevant to the performance of IOs. In addition, differentiating policy output along these dimensions enables a more fine grained measurement of IO performance.

By IOs, we mean formal intergovernmental, multilateral and bureaucratic organizational structures established to further cooperation among states (Martin and Simmons 2012; Rittberger et al. 2012). IOs consist of both interstate bodies (e.g., ministerial councils) and supranational bodies (e.g., secretariats), whose relative decision-making power is determined by institutional design and varies across IOs (Hooghe and Marks 2015). Observing the common distinction between organizations and institutions (North 1990: 5), the former refers to such formal organizational entities, and the latter to the principles and procedures that make up the rules of the game within IOs.

The performance of IOs merits attention because of its relevance to the effectiveness and legitimacy of IOs. First, scholarship suggests that IO performance is an important determinant of international regimes' problem-solving effectiveness (Miles et al. 2002; Young 2011). Second, research argues that IO performance influences the legitimacy of IOs (Buchanan and Keohane 2006; Dellmuth and Tallberg 2015). Unless IOs develop policy aimed at solving the transnational problems societies confront, they will fail to generate public support and be of limited value in combatting environmental degradation, intrastate conflict, financial instability, human-rights abuse, and other challenges.

Beyond the literature on IO performance, this article relates to two other bodies of research, neither of which has systematically explored the policy output of IOs. First, an extensive body of research in the field of International Relations (IR) examines the

effectiveness of international regimes. Generally, this research assesses if and when international regimes achieve their goals in terms of problem solving, focusing primarily on environmental issues (e.g., Young 1999; Miles et al. 2002). While international regimes in other issue areas have received growing attention, including human rights (Simmons 2009), peacekeeping (Fortna 2004), and economics and trade (Gray and Slapin 2012), there is to date limited comparative research on international regimes' effectiveness.

Second, there is an impressive literature in Comparative Politics (CP) on the performance of national political institutions. This literature includes research on the performance of democracies (Lijphart 1999; Schmidt 2002; Roller 2005), performance in discrete policy areas (Daugberg and Sønderskov 2012; Mackenbach and McKee 2013), and the performance of national institutions, such as parliaments (Arter 2006). While this literature focuses on policy outputs, there have been few systematic attempts to extend this analysis to the international level.

This article proceeds in three parts. First, we explain why we suggest an output-based approach to IO performance, discussing the merits and demerits of this conceptualization compared to alternatives, such as behavioral change and problem-solving effectiveness. Second, we present the five dimensions of policy output and explain how they individually and in combination allow for a measurement of IO performance. Third, we conclude by discussing how IO performance may be studied methodologically and explained theoretically.

Conceptualizing IO Performance: A Policy Output Approach

Existing research in CP and IR usually distinguishes between three measures of an institution's performance: its output, outcome, and impact (Young 2001; Roller 2005; Gutner and Thompson 2010). Output refers to the policies of an institution; outcome to the

4

implementation of these rules and programs through behavioral change by targeted actors; and impact to the contribution of these behavioral changes to the solution of the underlying problem.[1] For instance, the output of an environmental IO could be a resolution on the protection of endangered species, while outcome refers to the implementation of this decision, and impact concerns its effect on the species' population.

All three measures have strengths and weaknesses. We propose a policy output approach to IO performance for three reasons. First, policy output is causally prior to both outcomes and impacts, and therefore presents a logical and necessary first step in studying the performance of IOs. Policy output is seldom enough to conclude that IOs are successful problem solvers, but it is likely to be a requirement. In addition, policy output is necessary to consider even for studies focusing on alternative measures, as a way of verifying the contribution of IOs to behavioral change and identified impact. For example, studying implementation rates without considering the nature and volume of an IO's policy output can lead to a biased assessment, since low levels of output make it easier for IO member states to comply.

Second, output is more amenable to studying the causal influence of IOs. The causal chain linking an IO to output is shorter and less complex than the link between IOs, outcome, and impact. Consequently, outputs are results that can more clearly be attributed to an IO, while outcome and impact are more likely to be influenced by confounding factors. An outcome measure such as domestic compliance is more removed from the IO, making it more difficult to identify the IO's contribution. For example, low compliance does not necessarily mean that IOs have done a poor job. Domestic politics, bureaucratic incapacity, and economic turbulence can all contribute to a compliance gap. Likewise, making causal inferences about the impact of IOs is often challenging. IOs operate in complex environments, and impacts

---

[1] Note that parts of the CP literature uses the term outcome to refer to a policy's effect on the underlying problem.

might not be directly related to the actions of IOs. Even when faithfully implemented, ambitious and well-designed policies could have limited impact on the problems addressed by IOs because of external factors. Conversely, positive effects could result from factors unrelated to the efforts of IOs. A further complication in isolating the impact of individual IOs is overlaps in mandate between different organizations within a regime complex (Pollack and Hafner-Burton 2010: 289; Young 2011).

Third, the output measure facilitates comparative analysis of IO performance. Output can more readily be compared across and within IOs using standardized measures. As we will explain, this may be done based on five dimensions of policy output. Outcome and impact are more difficult to compare across IOs. For example, how does one compare the implementation of an environmental decision to that of a human rights decision? Compliance rates or impact indicators are notoriously difficult to aggregate and compare. Comparisons of outcome and impact are therefore mainly feasible with highly similar IOs, and generalizations across issue areas are difficult since most impact indicators are problem-specific.

Another challenge for outcome and impact measures is availability of comparable data. Most studies that apply impact measures use indicators from existing databases, like data on economic growth, unemployment, or the emission of pollutants. Statistics at the country level are nowadays comprehensive and reliable, although often limited to particular groups of countries (e.g., the OECD). Global indicators for impact performance are not widely available. In the absence of reliable and comparable indicators of outcome and impact, data on policy output offers a feasible way forward.

While these merits make the output-based approach to performance attractive, this measure comes with some drawbacks. While logically prior to outcome and impact, output does not necessarily matter for an IO's success in solving the problems it has been established to address. High policy performance can be thwarted by insufficient implementation and

external factors beyond the control of the IO. Moreover, even if policy output may be more amenable to comparison than outcome and impact, this measure, too, encounters problems in the comparative setting. Since there are no standardized ways of adopting and reporting policy in global governance, data on outputs vary in format and accessibility across IOs. Finally, a focus on the policy output of IOs in terms of rules and programs risks overlooking the informal side of global governance, in terms of non-codified institutions, practices, and norms. Yet, balancing these strengths and weaknesses, we see a significant and unexploited potential in advancing the study of IO performance by complementing existing outcome- and impact-oriented accounts with an output-based approach.

Both CP and IR offer illustrations of the usefulness of an output-based approach, even if it does not feature prominently in existing work. In CP, Putnam (1994), for instance, uses data on policies and expenditure to assess institutional performance in Italy. Other examples include Jahn and Müller-Rommel's (2010) analysis of social, environmental, and economic policy performance in Central and Eastern Europe, and Dasgupta et al.'s (2001) assessment of environmental policy performance in 31 countries. Studies of legislative performance commonly refer to the regulatory output of parliaments (Arter 2006). In IR, output-based studies of performance are rarer. One exception is Pollack and Hafner-Burton (2010), who analyze the EU's performance on gender mainstreaming based on policy outputs. Although many scholars study the outputs of IOs in general, and the EU and UN in particular (Golub 1999; Holloway and Tomlinson 1995), few link these analyses to questions of performance.

Five dimensions of IO Policy Output

Following our conceptualization of IO performance as the production of policy output that matches underlying cooperation problems, we propose a typology of policy output with five

features: volume, orientation, type, instrument, and target (Table 1). These features are applicable to the policy of all IOs, cover the most substantive aspects of output, and are of theoretical interest in the field. Differentiating policy output along these five dimensions enables a rich, comparable, and fine-grained picture of the activity and performance of IOs.

[Table 1. Five Dimensions of IO Policy Output]

The first two features – volume and orientation – are properties of the full body of output produced by an IO in a given time period. Together, they capture the level and issue focus of the policy adopted by an IO. Is the IO highly productive or deadlocked? Does the IO pursue a consistent or expanding policy agenda? The subsequent three features – type, instrument, and target – are properties of individual policy decisions, which aggregate into patterns at the level of IOs. What is the purpose of the policy – regulatory, distributive, or other? What is the nature of the policy instrument – hard or soft law? Who are the targets of the decision – states, private actors, or others?

In the following section, we provide definitions and examples of these five dimensions, illustrate how previous research has used them, and discuss options for operationalization. We further elaborate how this typology can be applied and argue that it mainly serves two purposes. First, it can be used to map descriptive patterns of IO policy output in a comprehensive way. Second, the typology provides a new toolkit for an assessment of IO performance. Each dimension highlights an important aspect of performance, but has shortcomings as a stand-alone measure. In our view, the ensemble of all five dimensions has the potential to advance our understanding of IO performance.

Policy Volume

The first dimension refers to the volume of output produced by an IO. Studying the scope and change of the policy output may answer a number of questions related to IO performance. Have individual IOs and IO bodies become more or less productive over time? If policy volume has changed, can it be explained by organizational reforms, new accessions, conflicts among member states, or other factors? Are there similar patterns in the policy volume across IOs that could be interpreted as a common trend toward higher or lower performance, or do we observe systematic differences across IOs active in different issue areas or world regions?

The volume of policy output is a common measure in studies of domestic legislative performance (Olson and Nonidez 1972; Arter 2006). Damgaard and Jensen (2006), for instance, compare the number of bills adopted by the Danish Folketing over time. There are also a number of studies that assess the policy volume of IOs. Pollack and Hafner-Burton (2010: 294), for instance, build an index of gender-related policy output based on the total number of activities carried out by the European Commission. Similarly, scholars have assessed the policy volume of the UN General Assembly (Holloway and Tomlinson 1995) and the Security Council (Allen and Yuen 2014). In studies of international courts, the volume of decisions often features as an indicator of their effectiveness (e.g., Alter 2014).

Policy volume is typically operationalized in quantitative terms, by the number of decisions within a given period. Measuring the quantity of an IO's policy output allows for a definition of zero – when an IO has been completely inactive. It also gives a precise representation of changes in an IO's productivity over time. Empirically, the yearly volume of policy output varies greatly. Some IOs only issue single declarations or a small set of decisions as the result of an annual summit, as in the case of the North Atlantic Council of the North Atlantic Treaty Organization (NATO) or the IGAD Assembly of Heads of State and Government of the Intergovernmental Authority on Development (IGAD). Other IOs produce

dozens, or even hundreds of decisions per year, like the World Health Assembly of the World Health Organization (WHO) and the EU Council of Ministers. One possible explanation for the lack of comparisons across IOs in existing research is the widely varying types and volumes of policy output.

We argue that there are three principal ways to address these difficulties. First, policy volume must be weighted. Potential weighting factors include text length, the scope of the IO's mandate, and the annual number of meetings of an IO's decision-making body. Second, by developing benchmarks for the policy volume of individual IOs, or highly similar organizations, it is possible to compare output across IOs in terms of relative changes of their volume (Pollack and Hafner-Burton 2010: 294). Mitchell (2004: 133f) discusses two operationalization strategies to facilitate comparison across international regimes. His first alternative is similar to developing benchmarks, and transforms data to an index based on a selected year that is set at 100. His other alternative is to use first differences that indicate annual changes in absolute levels.

Policy Orientation

A second principal dimension of an IO's overall policy output is its agenda orientation, in terms of the policy issues covered by IO decision-making. Like volume, this feature applies to the total corpus of an IO's output. A convenient way of characterizing the agenda orientation of IOs is according to the proportions of output devoted to different issue areas, such as development, security, and human rights.

The policy orientation of IOs draws attention to a range of important issues. First and foremost is the responsiveness of an IO to changing societal problems, as revealed by its policy orientation. In CP, scholars have studied the evolution of domestic policy agendas, demonstrating how they often develop incrementally for long periods, only to radically shift

when some issue rouses political attention (Baumgartner et al 2006). Do we see similar patterns among IOs? Scholars have begun to grapple with these questions (Alexandrova et al 2012), but remain focused on a limited set of IOs (mainly the EU) and have not yet presented frameworks for comparative analysis.

When policy agendas change over time, a second set of questions arises about the sources and patterns of such change.. Are some IOs typically the leaders of new policy agendas and others followers? For example, the World Bank frequently has pioneered policy, then mimicked by regional development banks. Are IOs sensitive to policy change in other organizations within a regime complex, as diffusion theories might expect? Further, are IOs with particular institutional rules (e.g., low hurdles), membership composition (e.g., democracies), or positions within regime complexes (e.g., more exposed to competition), more likely to adapt to new policy needs and trends? Institutional theory often suggests that IO policies are prone to stability, inertia, and resistance to reforms, as a result of institutional hurdles, organizational cultures, and adaptation costs.

The central task in operationalizing policy orientation is assigning policy topics to policy acts. Such classification instruments can proceed based on pre-defined topic lists, such as those used in the Policy Agendas Project (Baumgartner and Jones 1993), or by inductively building topic lists for each IO. The former has advantages for comparability, but if capturing nuanced variation for all types of IOs is the priority, the latter is preferable. Beyond dictionary-based methods, we envision that machine learning would be a useful tool, especially for classifying output from high-output IOs, since greater volume increases the reliability of automatized approaches (Grimmer and Stewart 2013).

Policy Type

The third dimension of IO output is policy type. Building on studies of domestic political systems (Almond and Powell 1966; Lowi 1972), we differentiate between regulatory, distributive, declarative, constitutional, and administrative output. Distinguishing between policy types may help to understand if and when IOs pursue their tasks through prescribing behavior, distributing goods, or setting common goals. Similarly, policy type can reveal the extent to which IOs concentrate on internally-oriented output (administrative and constitutional output) versus externally-oriented output (regulatory, distributive, and declarative output).

Regulatory output specifies actions that targeted actors are either expected to take or refrain from, aiming to achieve desired interactions by addressing problems of coordination and collaboration (Rittberger et al 2012: 120-123; Young 1999: 26-27). Regulatory output that addresses coordination problems could include rules on the use of sea lanes and common standards on telecommunication. International regulation can also address collaboration problems, such as those concerning nuclear proliferation or climate change.

IOs can also produce (re)distributive output to structure how goods and services are to be distributed among actors. While distributive output provides for distribution in a way that generates benefits for some without significant costs to others, such as the allocation of Internet domains, redistributive output implies that some actors benefit while others incur costs, such as the resource transfers in many World Bank programs.

IO output can also be declarative, asserting a joint position of member states or the IO. Declarative output may be symbolic, but can also assert a new agenda for an IO, lay out common goals, or promote actions while condemning others. For example, such policy would include UN General Assembly resolutions recognizing the 70th anniversary of the end of World War II, as well condemnations by EU member states of the actions of the Islamic State in Syria and Iraq.

Another type of policy output is constitutional, which concerns the rules governing the IO. [2] For instance, IOs may change their organizational structure or provide for new bodies or processes. Examples of constitutional output include the adoption of the WTO's Dispute Settlement Understanding, which created the Appellate Body, and the UN Convention against Torture, which established the Committee against Torture.

The last type of policy output is administrative. This type of policy output concerns internal operations and management of an IO. Administrative policy relates to the staffing and budgets of IOs, such the resolution adopted by the Organization of American States (OAS) in 1994 to modify the compensation system of its General Secretariat. It may also concern IO infrastructure, including buildings and facilities, such as the 2006 decision by the UN General Assembly to finance the so-called Capital Master Plan to renovate the UN's New York Headquarters.

Policy type could be operationalized as a categorical variable, with each type corresponding to one category. Such a variable would allow for the aggregation and analysis of temporal patterns within any IO. Coding could potentially rely on proxies, such as the mandate of the particular body whose output is being examined or the policy issue area, but content analysis of output is likely to yield more valid results. One approach is to identify a set of keywords and syntactic structures as indicators of each policy type, similar to the approach adopted by Koremenos (2013) in her coding of cooperation problems underlying international treaties. Any such coding instrument will face tradeoffs between specificity and comparability, so its development should harmonize with the particular research question at hand. If it centers on cross-IO analysis, comparability should be privileged; if the interest is temporal change within an IO, specificity is more important.

---

[2] Constitutional output does not include original founding treaties because these are not technically products of the IO.

Policy Instrument

A fourth feature of an IO's output are the instruments used to adopt individual policy decisions. Policies may be adopted through binding provisions (hard law) or non-binding provisions (soft law). Identifying the applicable policy instruments allows for an evaluation of when and where IOs opt for hard policies with a high level of bindingness or rely on soft recommendations.

The distinction between hard and soft law has attracted attention in IR and international law scholarship in recent years (Abbott and Snidal 2000; Guzman and Meyer 2009; Shaffer and Pollack 2013). At the center of this literature are three sets of questions (Shaffer and Pollack 2013). First, why and under what conditions might states opt to conclude agreements of a hard or a soft nature, and what advantages and disadvantages do hard and soft law present to states from an ex ante negotiating perspective? Second, how do hard and soft law interact? Are they complementary or competing policy instruments? Third, what are the consequences of hard and soft law for compliance with and effectiveness of policy?

Most IOs use a combination of hard and soft policy instruments. In the EU, for instance, the dominant forms of legislation are directives and regulations, which are both binding on their targets. However, EU policy in domains such as foreign policy and employment policy mainly consists of non-binding joint positions or goals to which the member states commit. Likewise, the UN combines binding resolutions in the Security Council with non-binding resolutions and declarations in the General Assembly.

The relation between hard and soft law may be seen as a dichotomy or as a continuum. In the legal positivist tradition, it is common to think in terms of a binary binding/non-binding divide. In this perspective, what makes law distinctive is its binding character – its capacity to impose legal obligations. Soft law, then, is the exception from law in the conventional sense.

Work in the rational choice tradition, by contrast, tends to conceive of the relation between hard and soft law as a continuum. In this vein, Abbott and Snidal (2000) conceptualize hard and soft law along a continuum of legalization, defined by the precision of rules, level of obligation, and delegation of interpretation and implementation to third parties. Law is hard when it scores high on precision, obligation, and delegation, and soft as it is weakened along these dimensions.

Whereas the simplest approach to operationalize policy type consists of a binary variable that distinguishes between binding and non-binding policies, a more complex measure based on Abbott and Snidal's (2000) dimensions could go beyond mere categorization and allow for a multi-dimensional ordinal ranking. Irrespective of approach, bindingness could be coded based on proxies, such as the mandate of the issuing IO body, or on close analysis of policy provisions. For example, some scholars use 'entry-into-force' as textual marker of binding policies (Mitchell 2003), whereas others (Hooghe and Marks 2015) rely on indicators such as the existence of opt-out clauses and annulment options.

Policy Target

The fifth dimension of IO output is the policy target, the actor or set of actors who is on the receiving end of the policy output. The targets are the individuals, organizations, communities or states whose behavior is prescribed, who are to receive benefits, who are expected to achieve results, or who are to take actions (Ingram and Schneider 1991). Like domestic policy, international policy output can target different actors, and in developing policy IOs have to choose who is to be targeted. Isolating the targets of policy makes it possible to evaluate whether IOs increasingly aim their actions at societal actors directly, rather than

work through governments, as implied in the conventional model of international public law (Moravcsik 2013).

IO policy output is likely to have one or more of the following targets: states, the IO itself, other IOs, nonstate actors, or the public or international community. The most traditional targets of IO policy are states. States that are targeted may include both member states and non-member states. For instance, while the EU generally targets its member states, it has increasingly addressed non-member states as it has moved toward a common foreign policy, as exemplified by the sanctions against Russia for the annexation of the Ukraine.

IO policy output may also target the IO itself, as is the case of constitutional and administrative output. For example, output that specifies the rules of procedure for an IO body intends to regulate its own functioning. Likewise, IO output concerning budget or staffing has the organization itself as the policy target. However, there are examples of other types of output that also target the IO itself. For instance, World Bank policies on environmental assessment, indigenous peoples, and involvement of NGOs are all policy output that set criteria for how the Bank's programs must be conducted.

In addition, an IO may target other IOs. Two examples are the UN Convention against Transnational Organized Crime and the UN Convention against Corruption, targeting both states and regional economic organizations. Other policy output aimed at IOs is more declarative, such as Resolution 68/262 of the UN General Assembly, which called upon states, IOs and other specialized agencies to not recognize the annexation of Crimea. Similarly, the African Union (AU) has called upon the UN Security Council to withdraw the referral of Omar El Bashir to the International Criminal Court (ICC) and on the ICC to postpone the trial of Uhuru Kenyatta.

Non-state actors, including individuals, non-profit, and for-profit organizations, may also be the target of IO policy. International criminal law, for instance, usually proscribes

behavior and attributes criminal negligence to individuals. Multinational corporations are the target of the Guidelines for Multinational Enterprises and Principles of Corporate Governance of the Organization for Economic Co-operation and Development (OECD). Similarly, the 1995 Code of Conduct for Responsible Fisheries of the Food and Agriculture Organization (FAO) targets states, IOs and non-state actors. A final example is the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES), which aims to stop individuals from smuggling and trading in endangered species.

Finally, a policy target of IO output may be the international community or the public. The UN General Assembly at times has targeted the public in its resolutions. For instance, in 1981, it declared the third Tuesday in September as the International Day of Peace, inviting "…non-governmental organizations, peoples and individuals to commemorate…and to co-operate with the United Nations in the observance of that Day" (A/RES/36/67).

This dimension may be operationalized through a categorical variable that classifies the policy target of IO output according to the three sets of actors identified above. A more fine-grained alternative leads to six different categories, beginning with a basic distinction between internal and external policy target, and in a second step distinguishing three groups of political actors: IOs, states, and non-state actors. Internally, the target can be the IO itself, its own member states or affiliated non-state actors. External policy targets are other IOs, other states, and non-state actors without formal linkages to the IO.

Using the Typology to Study the Output and Performance of IOs

The typology may be used for two primary purposes: to descriptively map the output of IOs on theoretically interesting dimensions, and to assess the performance of IOs. To begin, the five dimensions of volume, orientation, type, instrument, and target allow for a disaggregated

perspective on the structure of an IO's policy output. The five-fold typology spans a large number of policy configurations, which can be exhaustively mapped when combined with suitable indicators. This typology thus facilitates descriptive analysis of comparative patterns in IO performance as well as provides dependent variables for explanatory accounts of IO performance.

First, it provides researchers with a tool to identify temporal trends and patterns: Is the volume of output increasing or decreasing? Do we see shifts in the agenda orientation toward certain topics? Does IO policy increasingly address itself to non-state actors? Second, the five dimensions can be used to map variation across IOs: Do some IOs produce more output than others? Are some IOs more likely to employ hard law, or address enforcement problems? Third, identifying patterns of policy output across the five dimensions opens up the possibility of distinguishing different categories of IO output. Are some policy configurations more probable than others? We anticipate that patterns of co-variation will emerge in empirical data. For example, some output *types* are likely to be associated with certain output *instruments*: declarative policy is almost certain to be soft law, whereas regulatory or administrative policies are likely to be hard law. Both of the UN resolutions mentioned in Table 1, among others, displayed such associations between output dimensions. Likewise, some output *types* and *instruments* will likely correlate with specific *targets*: administrative output will tend to be directed toward the IO itself, and hard law output will primarily target states, the principal subjects of international law. Identifying such patterns of policy output across the five dimensions opens up the possibility of categorizing IOs. For instance, some IOs may be "talk shops," producing high volumes of declarative, non-binding policies with low target specificity.

In addition, the proposed typology could provide descriptive patterns related to explanatory questions of IO performance. There is an intuitive link between some output

dimensions and performance. For instance, policy volume may be seen as a measure of IO productivity, which is then used as a baseline for performance: an IO that fails to produce any policy output is clearly not performing. Given a constant mandate and problem structure, data on policy volume over time may indicate if an IO becomes more or less productive. Similarly, policy orientation would allow researchers to assess an IO's responsiveness to the problems and issues that emerge within its mandate. All other things equal, an IO that responds to all or most emerging issues could be said to perform better than an IO responding to only some or none.

For the remaining three dimensions, performance is best evaluated as the degree of correspondence between output and the problems facing an IO. This involves matching the structure of a policy, mapped along the typological dimensions, to the structure of the underlying cooperation problem at hand. Since the proposed typology's features correspond to the components of cooperation problems as they are commonly understood (Snidal 1985), it allows for the evaluation of how well policy matches up. From a performance perspective, *policy type* would then be evaluated according to whether it correctly recognizes and reflects the nature of the cooperation problem at the heart of an issue. Typically, distribution problems require distributive provisions, collaboration problems require regulatory provisions, and coordination problems require provisions that enable actors to make mutually consistent decisions. As for *policy instrument*, the concern is whether the policy at hand provides adequate measures to address the disparate preferences of actors given the cooperation problem. Distribution and collaboration problems are better served by binding provisions, whereas non-binding provision will general work for coordination problems due to the self-enforcing qualities of coordination equilibria. For *policy target*, performance is likely to depend on whether a policy directs itself to the range of actors involved in the cooperation problem at hand.

On the basis of our typology, well-performing IOs formulate and adopt policy provisions that appropriately respond to the full range of issues within their mandate, accurately targeting the underlying cooperation problems. Our typology provides a comprehensive tool with which researchers can avoid the limitations inherent in any one-dimensional approach to IO performance. Taken on their own, none of the five dimensions provides an adequate picture of IO performance. Policy volume, for instance, has limitations as a stand-alone measure of performance. An IO that produces a lot of output, but fails to target the responsible actors or use the appropriate instruments, certainly has a performance deficit. Likewise, an IO that formulates policies that match the underlying problem structure, but does not succeed in adopting such policies in adequate quantities, will also have a performance deficit. Only if we apply the five dimensions of output as an ensemble can we get a full and nuanced perspective on IO performance.

Conclusion: Methods and Theory in the Study of IO Performance

To conclude, we broaden the perspective to consider methods and theory in the study of IO performance. Given an output-based conceptualization, what methods and theories are most useful in mapping and explaining performance across a broad range of IOs?

Methodologically, the output-based approach we propose was borne out of a conviction that comparative research offers the way toward a more advanced understanding of IO performance. For this purpose, a conceptualization of performance that is conducive to comparison was required. As the research agenda on IO performance develops, we expect this approach to reveal spatial and temporal variation in IO performance. We also expect an output-based approach to be helpful for both qualitative and quantitative comparative studies.

Given the respective advantages of qualitative and quantitative studies, mixed-method designs are likely to be particularly rewarding.

Large-N analysis makes it possible to identify descriptive patterns in IO performance across time, IOs, issue areas, and world regions. Large-N analysis also offers the possibility of assessing the explanatory power of a broad range of factors. Since we do not expect mono-causal accounts of comparative patterns, this is a central strength. As suggested earlier, we expect automated text analysis to be a useful tool for classifying and measuring output along several dimensions. Case studies offer particular advantages for studying IO performance in context. Explanatory accounts of IO performance can benefit from in-depth case analysis of specific IOs, tracing causal mechanisms in the decision-making process and identifying how explanatory factors enhance or hinder output. Moreover, case studies are particularly useful when seeking to understand how the five dimensions of policy output interact in expressing the performance of IOs.

Theoretically, we expect accounts privileging power distribution, preference homogeneity, institutional design, and organizational culture to offer explanatory leverage. While sometimes formulated as competing explanations, we expect that these factors complement each other in accounting for IO performance. These explanations are anchored in classic approaches to international cooperation, and span factors both internal and external to IOs. At the same time, we do not expect these factors to exhaust the potentially relevant explanations (see, e.g., Gutner and Thompson 2010).

First, the distribution of power among the member states of an IO may affect performance. As suggested by hegemonic stability theory, dominance by a single state may be conducive to IO performance. When one state enjoys widely superior resources, it is well-positioned to buy the support of other countries through side payments (Drezner 2007; Stone 2011). Compensating losers who would otherwise block a policy, the dominant state ensures

passage of the policy it favors. Germany's role as Europe's paymaster, underwriting ambitious new EU policy programs, is a case in point, as is the United States' extension of financial favors to members of the UN Security Council in return for political support (Vreeland and Dreher 2014).

Second, the distribution of state preferences may shape IO performance. Where preferences are relatively homogenous, states are more likely to agree on new IO policy. Conversely, when state preferences diverge extensively, coming to any agreement may be exceedingly difficult. While the capacity of EU Council of Ministers to adopt policy is an example of the first, the UN Security Council's inability to respond adequately to conflicts during the Cold War is an example of the latter. While here treated as exogenous, state preferences may be a product of domestic factors such as economic interest, political regime, and national identity.

Third, the institutional design of IOs may impact their performance. Given a particular distribution of preferences, lax decision rules (qualified or simple majority) should make it easier for an IO to adopt policy than demanding decision rules (consensus or unanimity) (Scharpf 1997). Institutional access for private actors may improve the information that goes into decisions, making them more targeted, but also slow down the speed of policy-making (Tallberg et al. 2013). Autonomy for international bureaucracies, mandated to advance the goals of the organization, may make IOs more able to develop relevant policy (Hawkins et al. 2006). Agenda-setting rules, by privileging one or multiple agenda setters, may impact the efficiency and stability of policy adoption (Fiorina and Shepsle 1989).

Fourth, the organizational culture of an IO may influence its performance. IOs may fall prey to dysfunctional behavior because of internal cultures that reward other actions than those favoring targeted policy responses (Barnett and Finnemore 2004). Related, IOs may be organizationally structured in ways that make policy difficult to develop (Hanrieder 2015).

Comparative research on the crisis responses of IOs suggests that variation in organizational culture is a prominent explanation (Hardt 2014). A recent example is the WHO, whose late and inadequate policy response to the Ebola epidemic in Western Africa has been blamed on the absence of an emergency operations culture within the organization.

Bibliographical note:

Jonas Tallberg is Professor at the Department of Political Science, Stockholm University. Thomas Sommerer is Associate Professor at the Department of Political Science, Stockholm University. Theresa Squatrito is Post-doctoral Research Fellow at PluriCourts Center of Excellence, University of Oslo. Magnus Lundgren is Post-doctoral Research Fellow at the Department of Political Science, Stockholm University.

Address for correspondence: Jonas Tallberg, Department of Political Science, Stockholm University, SE-106 91 Stockholm, Sweden. Email: jonas.tallberg@statsvet.su.se

References

Abbott, K.W. and Snidal, D. (2000) 'Hard and soft law in international governance',
International Organization 54(3): 421–56.

Alexandrova, P., Carammia, M, and Timmermans, A (2012) 'Policy punctuation and issue
diversity on the European Council Agenda', *Policy Studies Journal*, 40(1): 69–88.

Alter, K. (2014) *The new terrain of international law: courts, politics, rights*, Princeton:
Princeton University Press.

Allen, S.H. and Yuen, A.T. (2014) 'The politics of peacekeeping: UN security council
oversight across peacekeeping missions', *International Studies Quarterly* 58(3): 621-32.

Almond, G.A. and Powell, G.B. (1978) *Comparative Politics: System, Process, and Policy*,
Boston: Little, Brown.

Arter, D. (2006) 'Introduction: Comparing the legislative performance of legislatures',
Journal of Legislative Studies 12(3-4): 245–57.

Barnett, M. and Finnemore, M. (2004) *Rules for the World: International Organizations in
Global Politics*, Ithaca: Cornell University Press.

Baumgartner, F. R., Green-Pedersen, C., and Jones, B. D. (2006) 'Comparative studies of
policy agendas', *Journal of European Public Policy*, 13(7): 959-974.

Baumgartner, F.R. and Jones, B.D. (1993) *Agendas and instability in American politics*.
University of Chicago Press.

Buchanan, A. and Keohane, R.O. (2006) 'The Legitimacy of Global Governance Institutions',
Ethics and International Affairs 20(4): 405-437.

Damgaard, E. and Jensen, H. (2006) 'Assessing strength and weakness in legislatures: The
case of Denmark', *Journal of Legislative Studies* 12(3-4): 426–42.

Dasgupta, S., Mody, A., Roy, S. and Wheeler, D. (2001) 'Environmental regulation and development: A cross-country empirical analysis', *Oxford Development Studies* 29(2): 173–87.

Daugbjerg, C. and Sønderskov, K.M. (2012) 'Environmental policy performance revisited: Designing effective policies for green markets', *Political Studies* 60(2): 399–418.

Dellmuth, L.M. and Tallberg, J. (2015) 'The social legitimacy of international organisations: Interest representation, institutional performance, and confidence extrapolation in the United Nations', *Review of International Studies* 41(3): 451-475.

Drezner, D. (2007) *All Politics is Global. Explaining International Regulatory Regimes*. Princeton, NJ: Princeton University Press.

Easton, D. (1965) *A Framework for Political Analysis*, Englewood Cliffs: Prentice Hall.

Elsig, M. (2010) 'The World Trade Organization at work: Performance in a member-driven milieu', *Review of International Organizations* 5(3): 345-363.

Falkner, G. (2005) *Complying with Europe: EU Harmonisation and Soft Law in the Member States*, Cambridge: Cambridge University Press.

Fiorina, M.P., & Shepsle, K.A. (1989) 'Formal theories of leadership: Agents, agenda setters, and entrepreneurs?', in B.D. Jones (ed.), *Leadership and Politics: New Perspectives in Political Science*. Lawrence, Kans.: University Press of Kansas, pp. 17-40.

Fortna, V.P. (2004) 'Does peacekeeping keep peace? International intervention and the duration of peace after civil war', *International Studies Quarterly* 48(2): 269–92.

Golub, J. (1999) 'In the shadow of the vote? Decision making in the European Community', *International Organization* 53(4): 733–64.

Gray, J. and Slapin, J.B. (2012) 'How effective are preferential trade agreements? Ask the experts', *Review of International Organizations* 7(3): 309–33.

Grimmer, J. and Stewart, B.M. (2013) 'Text as Data: The Promise and Pitfalls of Automatic

    Content Analysis Methods for Political Texts', *Political Analysis*, 21(3): 267–297

Gutner, T. and Thompson, A. (2010) 'The politics of IO performance: A framework', *Review*

    of International Organizations 5(3): 227–48.

Guzman, A.T. and Meyer, T.L. (2010) 'International soft law', *Journal of Legal Analysis*

    2(1): 171–225.

Hanrieder, T. (2015) *International Organization in Time: Fragmentation and*

    Reform, Oxford: Oxford University Press.

Hardt, H. (2014). *Time to React: The Efficiency of International Organizations in Crisis*

    Response. Oxford University Press.

Hawkins, D.G., Lake, D.A., Nielson, D.L. and Tierney, M.J. (2006) *Delegation and agency in*

    international organizations, Cambridge: Cambridge University Press.

Holloway, S.K. and Tomlinson, R. (1995) 'The new world order and the general assembly:

    Bloc realignment at the UN in the post-cold war World', *Canadian Journal of Political*

    Science 28(2): 227–54.

Hooghe, L. and Marks, G. (2015) 'Delegation and pooling in international organizations',

    Review of International Organizations 10(3): 305-328.

Ingram, H. and Schneider, A. (1991) 'The choice of target populations', *Administration &*

    Society 23(3): 333–56.

Jahn, D. and Müller-Rommel, F. (2010) 'Political institutions and policy performance: A

    comparative analysis of Central and Eastern Europe', *Journal of Public Policy* 30(1):

    23–44.

Koremenos, B. (2013) 'The continent of international law', *Journal of Conflict Resolution*

    57(4): 653–681.

Lijphart, A. (1999) *Patterns of Democracy: Government Forms and Performance in thirty-six Countries*, New Haven: Yale University Press.

Lipson, M. (2010) 'Performance under ambiguity: international organization performance in UN peacekeeping', *Review of International Organizations* 5(3): 249-284.

Lowi, T.J. (1972) 'Four systems of policy, politics, and choice', *Public Administration Review* 32(4): 298–310.

Mackenbach, J.P. and McKee, M. (2013) 'A comparative analysis of health policy performance in 43 European countries', *European Journal of Public Health* 23(2): 195–201.

Martin, L. A. and Simmons, B. A. (2013) 'International Organizations and Institutions', in W. Carlsnaes, T. Risse, and B.A. Simmons (eds.), *Handbook of International Relations*, London: SAGE.

Miles, E., A. Underdal, S. Andresen, J. Wettestad, J. B. Skjaerseth and E. M. Carlin (2002) *Environmental Regime Effectiveness: Confronting Theory with Evidence*, Cambridge, MA: MIT Press.

Mitchell, R. (2003) 'International environmental agreements: a survey of their features, formation, and effects', *Annual Review of Environment and Resources*, 28(1), 429-461.

Mitchell, R. (2004) 'A Quantitative Approach to Evaluating International Environmental Regimes', in A. Underdal and O. Young (eds.), *Regime Consequences: Methodological Challenges and Research Strategies*. Amsterdam: Kluwer Academic Publishers, pp. 121-149.

Moravcsik, A. (2013) 'Liberal Theories of International Law', in M. A. Pollack and J. Dunoff (eds.), *Interdisciplinary Perspectives on International Law and International Relations*, Cambridge: Cambridge University Press, pp. 83–118.,

Oberthür, S., Jørgensen, K.E. and Shahin, J. (eds.) (2013) *Performance of EU in International Institutions*. London: Routledge.

Olson, D.M. and Nonidez, C.T. (1972) 'Measures of legislative performance in the US House of Representatives', *Midwest Journal of Political Science* 16(2): 269–77.

Pollack, M.A. and Hafner-Burton, E.M. (2010) 'Mainstreaming international governance: The environment, gender, and IO performance in the European Union', *Review of International Organizations* 5(3): 285–313.

Putnam, R.D., Leonardi, R. and Nanetti, R.Y. (1994) *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, NJ: Princeton University Press.

Rittberger, V., Zangl, B. and Kruck, A. (2006) *International Organization*, 1st edition, Basingstoke: Palgrave Macmillan.

Roller, E. (2005) *The Performance of Democracies: Political Institutions and Public Policy*, Oxford: Oxford University Press.

Scharpf, F.W. (1997) *Games Real Actors Play*. Boulder, Colo.L Westview Press.

Schmidt, M.G. (2002) 'Political performance and types of democracy: Findings from comparative studies', *European Journal of Political Research* 41(1): 147–63.

Shaffer, G. and Pollack, M.A. (2013) 'Hard and soft law', in M. A. Pollack and J. Dunoff (eds.), *Interdisciplinary Perspectives on International Law and International Relations*, Cambridge: Cambridge University Press, pp. 197-222.

Snidal, D. (1985) 'Coordination versus prisoners' dilemma: Implications for international cooperation and regimes', *American Political Science Review* 79(04): 923–942.

Simmons, B.A. (2009) *Mobilizing for human rights: international law in domestic politics*, Cambridge: Cambridge University Press.

Stone, R. W. (2011) *Controlling institutions: International organizations and the global economy*. Cambridge University Press.

Tallberg, J., Sommerer, T., Squatrito, T. and Jönsson, C. (2013) *The opening up of international organizations: transnational access in global governance*, Cambridge: Cambridge University Press.

Underdal, A. (2002) 'One question, two answers', in Miles, E., A. Underdal, S. Andresen, J. Wettestad, J.B. Skjaerseth, and E.M. Carlin (eds.), *Environmental Regime Effectiveness: Confronting Theory with Evidence, Cambridge,* MA: MIT Press*, pp. 3–45.

Vreeland, J. and Dreher, A. (2014) *The political economy of the United Nations Security Council: money and influence*, Cambridge: Cambridge University Press.

Weaver, C. (2010) 'The politics of performance evaluation: independent evaluation at the International Monetary Fund', *Review of International Organizations* 5(3): 365–385.

Young, O.R. (1999) *The Effectiveness of International Environmental Regimes: Causal Connections and Behavioral Mechanisms*. Cambridge, MA: MIT Press.

Young, O.R. (2001) Inferences and indices: evaluating the effectiveness of international environmental regimes', *Global Environmental Politics* 1 (1): 99-121.

Young, O.R. (2011) 'Effectiveness of international environmental regimes: Existing knowledge, cutting-edge themes, and research strategies', *Proceedings of the National Academy of Sciences* 108, 19853–60.

Table 1: Five dimensions of IO policy output

Output dimension	Definition	Level of analysis	Examples
Volume	The volume of policy decisions produced by an IO in a given time period.	IO, IO year	The UN General Assembly adopted 286 resolutions in its 69th (2014/15) session.
Orientation	The distribution of policy acts across issue areas in a given time period.	IO, IO year	The principal issue areas of UN General Assembly resolutions adopted in 2014/15 were security (28%), human rights (24%), and development (16%).
Type	Whether the aim of a policy is regulatory, distributive, declarative, constitutional or administrative.	Individual policy act	UN General Assembly Resolution 69/285, recognizing the permanent neutrality of Turkmenistan, is declarative. Resolution 69/258, financing the UN mission in South Sudan, is administrative.
Instrument	Whether a policy is binding (hard law) or non-binding (soft law).	Individual policy act	UN General Assembly resolution 69/285 contains non-binding recommendations, so classifies as soft law. Resolution 69/258 requires member states to contribute funds, imposing legal obligations, so has a higher level of bindingness.
Target	Whether a policy is aimed at states, nonstate actors, the IO itself, other IOs, or the international community as a whole.	Individual policy act	UN General Assembly resolution 69/285 targets states, nonstate actors, and the international community as a whole. Resolution 69/258 targets member states and the IO itself (UN Secretariat).