The Performance of International Organizations: An Output-Based Approach

Jonas Tallberg
Stockholm University
jonas.tallberg@statsvet.su.se

Thomas Sommerer
Stockholm University
thomas.sommerer@statsvet.su.se

Theresa Squatrito
University of Oslo
theresa.squatrito@jus.uio.no

Magnus Lundgren
Stockholm University
magnus.lundgren@statsvet.su.se

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. As an initial step toward explaining variation in IO performance, this paper discusses how performance may best be conceptualized for empirical analysis. We advance an output-based account of IO performance, privileging the results IOs produce in terms of policy output. This conceptualization is distinct from the main alternative perspectives on IO performance, emphasizing either behavioral change by targets (policy outcome) or problem-solving effectiveness (policy impact). In addition, we introduce a typology of output that captures five generic features of IO policy of theoretical importance: policy volume, orientation, type, instrument, and target. We conclude by highlighting the implications of an output-based account to IO performance for the study of institutional design and effectiveness in global governance.

Paper prepared for presentation at the General Conference of the ECPR,
Montreal, August 26-29, 2015.

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, adopting, and enforcing policy, others are less successful. Whereas the European Court of Human Rights is famous for its strong record, for example, the UN Human Rights Commission was abolished because of its poor performance. While the EU Council of Ministers 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?

As an initial step toward explaining variation in IO performance, this paper discusses how to best conceptualize performance for empirical analysis. We advance an output-based account of IO performance, privileging the results IOs produce in terms of policy output. Specifically, we argue that IO performance is best conceptualized as the capacity of IO bodies to produce policy output in terms of rules, policies, and programs in a given area (cf. Easton 1965). In contrast to alternative approaches, our conceptualization of IO performance is especially beneficial for studying the consequences of institutional design for IO accomplishments and for comparing the performance of IOs. To date, scholarship has been unable to systematically examine the link between the institutional design of IOs, such as the autonomy of IO bureaucracies and the rules governing decision-making, and the performance of IOs. Additionally, scholarship has made little headway on comparative analysis of IO performance. We contend that an output-based approach to IO performance is advantageous for addressing both these gaps in existing literature.

In developing this output-based account to IO performance, we introduce a taxonomy of policy output with five generic dimensions: policy volume, policy orientation, policy type, policy instrument, and policy target. Each dimension represents an issue of theoretical and empirical contention in existing scholarship. In addition, differentiating policy output along five dimensions enables a more fine-grained measurement, with added benefits for identifying patterns of variation in IO performance. For this reason, we also argue that each of these dimensions of IO policy output is critical to any overall assessment and characterization of IO performance.

Our conceptualization of IO performance builds on and contributes to three bodies of research, none of which has systematically and comparatively 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; Hafner-Burton 2013), peacekeeping (Fortna 2004), and economics and trade (Mansfield and Reinhardt 2008; Gray and Slapin 2012), to date, there is limited comparative research on the effectiveness of international regimes.

Second, there is an impressive literature in Comparative Politics (CP) on the performance of national political institutions. Contributions to this literature include research that focuses on: the performance of democracies (Lijphart 1999; Schmidt 2002; Roller 2005); policy performance in discrete issue areas, such as environmental policy (Daugberg and Sønderskov 2012), health policy (Mackenbach and McKee 2013), and economic policy (Alvarez et al. 1991); and the performance of different types of national institutions, such as parliaments (for an overview, see Arter 2006). While CP literature focuses on policy outputs,

there have been few systematic attempts to build on this scholarship to analyze performance at the international level.

Third, there is a limited body of research on the performance of IOs. In a pioneering article, Gutner and Thompson (2010) introduce a multi-layered understanding of IO performance that involves both process- and result-oriented standards, including policy output. However, empirical work on policy output has so far remained limited to a small set of major IOs, notably, the EU (Pollack 1994; Golub 1999; Thomson et al. 2006; Oberthür et al. 2013; SIEPS 2015), the UN (Holloway and Tomlinson 1995), and international courts (Alter 2014; Carruba and Gabel 2014).

This paper proceeds in three parts. First, we discuss the concept of IO performance and explain why we suggest an output-based approach. We discuss the merits and demerits of this conceptualization compared to alternatives, such as behavioral change and problem-solving effectiveness. Second, we develop a taxonomy of five dimensions of policy output amenable to empirical measurement: policy volume, policy type, policy instrument, policy target, and policy orientation. Third, the paper highlights the implications of IO performance for studying the consequences of institutional design, such as the role of IO bureaucracies, as well as the effectiveness of IOs.

Conceptualizing IO Performance: An Output-based Approach

This paper aims to contribute to the study of IO performance by providing a conceptualization of performance that emphasizes the capacity of IO bodies to produce policy output. In this section, we clarify the difference between effectiveness and performance, discuss alternative conceptualizations of performance, and argue that an analysis of IO performance should begin with a focus on policy output.

Why Performance?

Existing research in IR and CP tends to conceptualize the achievements of political institutions as either effectiveness or performance. The literature on international regime effectiveness has made it more common to speak of effectiveness in IR. Only recently have IR scholars begun to frame their research in terms of performance (Gutner and Thompson 2010). Conversely, in CP, performance has traditionally been the favored theoretical concept in studies of the accomplishments of domestic political institutions.

While both effectiveness and performance are well-established and useful concepts, we favor performance for two principal reasons (cf. Gutner and Thompson 2010). First, performance is more actor-centered than effectiveness. As everyday language would imply, performance refers to an organization's activities or completion of a task. Since our primary focus is IOs, viewed as actors and not as systems of rules, performance is more suitable. In contrast, the primary focus of effectiveness is problem-solving, and whether systems of rules successfully contribute to the desired problem-solving. In other words, effectiveness does not require an actor orientation, and rather looks to whether and to what extent a particular goal is achieved, irrespective of the actors that shape that achievement.

Second, performance is not conceptually bound to assessing organizational accomplishments in relation to the problem-solving goals of an organization. In IR, effectiveness is understood as a concept that compares a regime's results to its intended goals in terms of problem solving (Underdal 2004; Young 2011). In contrast, performance does not necessarily require an assessment based on a comparison of accomplishments against the goal of the regime's creators. Rather, performance may be evaluated through comparisons of accomplishments within and across organizations based on other metrics, such as the

conversion of resources into results. In this sense, performance is more amenable to comparison across IOs that have very different goals or objectives.

Why an Output-Based Approach?

In practice, effectiveness and performance studies in IR and CP are operationalized as the output, outcome, or impact of political institutions and regimes (Underdal 2002: 6f). Typically output refers to the rules, policies, and programs of an institution; outcome to the implementation of these rules through behavioral change by targeted actors; and impact to the contribution of these rules and behavioral changes to the solution of the underlying problem.[1] For instance, the output of an environmental IO could consist of a resolution on the protection of an endangered species, while outcome refers to the implementation of this decision, and impact concerns the effect of this policy on the species' population.

We use an output-based approach to IO performance for three primary reasons. First, output, in comparison to outcome and impact, 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. Thus, output can more readily accommodate explanatory accounts of the variation in performance. In slightly different terms, IO outputs are results that can more clearly be attributed to an IO.

In contrast, outcome and impact are likely to be influenced by a wide range of factors, many of which are outside the direct control of an IO. While outcome-based assessments of performance are common in IR, domestic implementation is more removed from the source of policies – the IO – and therefore more exposed to the problem of confounding factors. For example, low compliance with international rules does not necessarily mean that IOs have

[1] Note, however, that there is considerable confusion about the term outcome, which is used differently in different literatures. In large parts of the CP literature, but also sometimes in IR, the effects of a policy on the underlying problem are referred to as outcome.

done a poor job. Domestic politics, bureaucratic capacity deficits, and economic turbulence can all contribute to a possible implementation gap.

Impact performance is even more problematic for making causal inferences about the achievements of IOs. IOs operate in complex environments. Ambitious and well-designed policies could be ineffective because of external factors beyond the control of IOs. Conversely, positive effects on problems addressed by IOs could very well result from political or non-political factors at national or sub-national levels. Good as well as poor impacts might not be directly related to regulatory measures of IOs (Gutner and Thompson 2010: 237). For example, assigning low scores to a human rights committee because violations continue might be unfair, if the committee is active in supporting victims, training observers, and shaming states.

A further complication in isolating the effects of an individual IO on impact is the multiple overlaps in mandate between different organizations making up a regime complex (Keohane and Victor 2011). This means that, in addition to the general methodological challenge of environmental factors that cannot be controlled by political actors, it is even more difficult to assign identified impacts to IOs in global governance (Pollack and Hafner-Burton 2010: 289; Young 2011). Thus, policy output is the most direct measure of IO performance, since decisions clearly are within the control of the IO and little exposed to external, confounding factors. If we are interested in assessing the effects on IO performance of factors such as institutional design, organizational culture, and power distribution, output is by all standards the most suitable measure.

Second, output is advantageous because it can facilitate comparative analysis. Output can be compared across IOs as well as within IOs. Standardized measures of policy output, such as policy volume, could be a possible comparative metric for assessing the relative performance of IOs. Outcome and impact are more difficult to compare. 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 impact performance are therefore only possible within sets of highly similar IOs or IO bodies, and generalizations across issue areas very difficult since most impact indicators are problem-specific.

Another challenge for outcome and impact measures is availability of comparable data. Most of the studies that apply impact-based measures of performance use indicators from existing databases, like macroeconomic data on growth and unemployment, or data on the emission of pollutants. Statistics at the country level are nowadays comprehensive and reliable, although often limited to particular regions or groups of countries (e.g., the OECD). Yet, the availability of global indicators for impact performance is much more limited. When reliable and comparable indicators of outcome and impact are absent, policy output may even be the only feasible way forward. While there is minimal previously composed data on IO policy output (with some exceptions), a focus on policy output thus holds more promise for comparative studies of IO performance.

Third, 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. Unless IOs are successful in producing policies, they will be unable to achieve behavioral change and contribute to the solution of societal problems. For example, studying implementation rates without considering the nature and volume of an IO's policy output can lead to a biased view, since low levels of output will make it easier for IO member states to comply, and vice versa. Similarly, high output does not necessarily translate into behavioral change and problem solving, as policies may be undermined by both insufficient implementation and external factors beyond the control of an IO (Gutner and Thompson

2010: 236). Generally, we contend that data on policy output are essential for verifying the causal link between an IO's activities, behavioral change, and identified impact. For these reasons, we propose that a policy-output approach can be useful for studying the relationship between institutional design and output, as well as inform further inquiries into outcome and impact.

Both IR and CP offer illustrations of the usefulness of an output-based approach, even if it does not feature prominently in existing work.[2] We draw upon some important works that rely on output measures. For example, Putnam (1994) is one of the better known examples, as he uses data on policies and expenditure to assess the institutional performance in different parts of Italy (see also Roller 2005: 32). Other examples of studies in CP using output indicators are 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 also commonly refer to the regulatory output of parliaments (Arter 2006). Output-based studies of performance are rarer in IR, but there are recent examples. For instance, Pollack and Hafner-Burton (2010) analyze the performance of the EU in gender mainstreaming based on policy outputs. Although many scholars study the outputs of IOs in general, and the EU and UN in particular, only few link these analyses to questions of performance and effectiveness (Alesina et al 2005; Golub 1999; Holloway and Tomlinson 1995).

Five dimensions of IO Policy Output

[2] Outcome- and impact-based approaches are more common in both literatures. For an overview of impact performance in CP, see Roller (2006). Studies of impact performance in IR are common in the fields of human rights (e.g., Simmons 2009), environmental protection (e.g., Wettestad 2002), and conflict studies (e.g., Fortna 2004). Outcome performance is commonly used in studies on the implementation of EU legislation (e.g., Falkner et al. 2005).

How, then, should the policy output of IOs be conceptualized and measured? We propose a typology of policy output with five features: policy volume, policy orientation, policy type, policy instrument, and policy target. These features, summarized in Table 1, 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 makes it possible to arrive at a rich, comparable, and fine-grained picture of the performance of IOs.

[Table 1. Five Dimensions of IO Policy Output]

The first two features – policy volume and policy 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 – policy type, policy instrument, and policy target – are properties of individual policy decisions, which aggregate into patterns at the level of IOs. What is the purpose of the intended policy – regulatory, distributive, or other? What is the nature of the policy instrument – hard law or soft law? Who are the targets of the decision – states, private actors, or others? We believe that each of these dimensions to IO policy output is central to any overall characterization of IO performance.

Policy Volume

The first dimension on which policy performance may be assessed 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 the performance of IOs. Have individual IOs and IO bodies become more or less productive over time? If the policy volume has changed, can this 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 and world regions?

The volume of policy output is a common measure in studies of legislative performance at the domestic level (e.g., 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. Alesina et al. (2005) have mapped the expansion of EU legislation by the number of legal acts (see also Pollack 1994; Golub 1999). Similarly, scholars have assessed patterns in the policy volume of the UN General Assembly (e.g., Marín-Bosch 1987; Holloway and Tomlinson 1995), and the UN Security Council (e.g., Allen and Yuen 2014). In studies of international courts, the volume of decisions often features as an indicator of their effectiveness (e.g., Cichowski 2006; Alter 2014).

The 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 absolute zero – when an IO has been completely inactive. It also gives a precise representation of changes in the productivity of an IO over time. Empirically, the yearly volume of policy output varies enormously. 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 APEC Economic Leaders' Meeting, the NATO North Atlantic Council, or the IGAD Assembly of Heads of State and Government. Other IOs produce dozens, or even hundreds of decisions per year, like the WHO World Health Assembly, the OECD Council, the UN Security Council, and the

EU Council of Ministers. One possible explanation for the absence of comparisons across IOs in existing research is likely to be the difficulties involved in comparing such widely varying types of policy output.

We argue that there are two principal ways to approach these difficulties. First, policy volume needs to be weighted. Potential weighting factors include text length, the scope of the IO's mandate, and the number of meetings of an IO's decision-making body, since some IOs only meet on an annual or biannual basis, whereas others convene several times a year. Second, by developing benchmarks for the policy volume of individual IOs, or groups of highly similar organizations, it would be possible to compare performance across IOs in terms of relative changes of their productivity (c.f., Pollack and Hafner-Burton 2010: 294).

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 policy volume, this feature applies to the total corpus of an IO's output. A convenient way of characterizing the agenda orientation of IOs is in terms of 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. Multiple strands of institutional theory highlight why IO policies should be characterized by stability, inertia, and resistance to reforms. Expanding the portfolio of policies usually requires high-level decisions involving significant institutional hurdles, like consensus among the contracting parties (Scharpf 1988). The organizational cultures of IOs often privileges stability over change (Barnett and Finnemore 2004). Like all rules, IO policies tend to become self-reinforcing, by structuring expectations, presenting adaptation

costs, supporting vested interests, and generating positive feed-back effects that serve to lock in the status quo (Fioretos 2011). At the same time, being responsive to societal concerns is central for IOs, both from a normative and a problem-solving perspective.

In those instances that policy agendas change over time, this raises a second set of questions about the sources and patterns of such change. Are IOs with some characteristics in terms of institutional rules (e.g., low hurdles), membership composition (e.g., democracies), and position within regime complexes (e.g., more exposed to competition), more likely to be adaptive to new policy needs and trends? Are IOs sensitive to policy change in other organizations within the issue area or regime complex, as diffusion theories might lead us to expect? Are some IOs typically leaders and others followers in adopting new policies? An example would be development banks, where the World Bank frequently has functioned as policy pioneer then mimicked by regional development banks, as in the introduction of "participatory development" as the new policy credo.

Policy Type

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

IOs' regulatory output specifies actions that target actors (states, individuals, organizations, etc.) are either expected to take or refrain from, aiming to achieve desired interactions or joint benefits by addressing problems of coordination and collaboration (Rittberger et al 2012: 120-123; Young 1999: 26-27). Examples of regulatory output that addresses coordination problems include rules on the use of sea lanes and common standards on telecommunication and measurement. International regulation can also address collaboration problems, such as those concerning nuclear proliferation, trade in endangered species, or climate change.

In addition to regulatory output, IOs can have the capacity to produce (re)distributive output pertaining to the distribution of goods and services 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 and others incur costs, as in the resource transfers involved in many World Bank programs.

IO output can also be declarative, asserting a joint position of the member states or the IO. Declarative output may at times be of a symbolic nature, 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 (A/RES/69/267), as well joint condemnations by the EU's member states of the actions of the Islamic State in Syria and Iraq. Declarative output may be particularly common in response to emergencies, because such decisions can be easier to agree on and do not necessarily require adoption through a normal decision-making process. For example, EU statements are often used in crisis situations, as in European Council statements on the refugees crossing the Mediterranean in the spring of 2015.

Another type of policy output is constitutional output, which concerns the rules governing the IO. This is output that specifies the design of the IO and the rules that regulate how it functions, such as decision rules and rules of procedure. For instance, IOs may change their organizational structure or provide for new bodies or processes within the IO. Examples of constitutional output include the adoption of the WTO's Dispute Settlement Understanding, which created the Appellate Body and its rules, the UN human rights treaties that established monitoring committees (e.g., Committee against Torture created by the Convention against Torture), or EU treaty revisions that changed the powers of the European Parliament (e.g., Treaty of Maastricht establishing the co-decision procedure).[3]

The last type of policy out is administrative. This type of policy output concerns internal operational and managerial concerns of an IO. Administrative policy decisions relate to the staffing and budgets of IOs, such the resolution adopted by the Organization of American States in 1994 to modify the compensation system of its General Secretariat (Resolution 1275 XXIV-O/94). 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, the renovation of the UN Headquarters in New York (Resolution 61/237).

Policy Instrument

A fourth feature of an IO's output are the instruments used to adopt individual policy decisions. Policies may either 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.

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

The distinction between hard law and soft law has attracted significant attention in IR and international law scholarship in recent years (Lipson 1991; 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 the policy adopted?

The relation between hard law and soft law may either 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. Compared to binding hard law, soft law may be defined as "a codified instrument that is publicized, issued through an institutionalized process, with the aim of exercising a form of authority or persuasion, even though the instrument is not formally legally binding" (Shaffer and Pollack 2013: 198). Work in the rational choice tradition, by contrast, tends to conceive of the relation between hard law and soft law as a continuum. In this vein, Abbott and Snidal (2000) conceptualize hard law and soft law along a continuum of legalization, defined by three dimensions (1) precision of rules, (2) level of obligation, and (3) delegation of interpretation and implementation to third parties. Law is hard when it scores high on precision, obligation, and delegation, and then increasingly soft as it is weakened along one or several of these dimensions.

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, both of which are binding on

their targets. At the same time, 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.

Policy Target

The fifth and final dimension of IO output is the policy target, the actors or set of actors who are 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 or sets of 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. Perhaps the most traditional targets for 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 usually the case of constitutional and administrative output. For example, IO 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, involuntary resettlement, 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 integration organizations.[4] Other policy output aimed at IOs is more declarative, such as Resolution 68/262 of the UN General Assembly on the integrity of the Ukraine, which "calls upon all States, international organizations and specialized agencies not to recognize any alteration of the status of the Autonomous Republic of Crimea and the city of Sevastopol." Another organization that targets other IOs is the African Union, which targeted both the UN and the International Criminal Court when calling upon the UN Security Council to withdraw the referral of Omar El Bashir to the ICC for indictment or 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 OECD's Guidelines for Multinational Enterprises and the OECD's Principles of Corporate Governance. Similarly, the FAO's 1995 Code of Conduct for Responsible Fisheries 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 ultimately is aimed at stopping individuals from smuggling and trading in endangered species.

Finally, a policy target of IO output may be the international community or the public in general. The UN General Assembly at times has targeted the public in its resolutions. For

[4] See UN Convention against Transnational Organized Crime Article 36.2 and UN Convention against Corruption Article 67.2.

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).

Patterns and Combinations

Taken together, the five dimensions of volume, orientation, type, instrument, and target allow for a disaggregated perspective on the structure of an IO's policy output. With five separate dimensions, the typology spans a large number of possible policy configurations, increasing our confidence that it permits for exhaustive mapping, when combined with suitable indicators. Naturally, some policy configurations are more probable than others, and we anticipate that patterns of co-variation will emerge in empirical data. For example, some output *types* are likely to associate with certain output *instruments*: declarative policy is almost certain to have soft law characteristics, whereas regulatory or administrative policies are more likely to be hard law. Both of the UN resolutions mentioned in Table 1, and several other aforementioned examples, 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. Declaratory output will often target the international community and public as large, as exemplified in the UN Declaration of Human Rights, which directs itself to "all peoples and all nations."

While such patterns of co-variation are more likely to emerge among the policy-specific dimensions of type, instrument, and target, they may also occur for the IO-specific dimensions of *volume* and *orientation*. For example, it is conceivable that IOs leaning toward

the production of declarative or administrative output display high output volumes, as these policy types expose member states to lower financial and political costs. Similarly, it is conceivable that IOs focusing on issue areas with aspirational qualities (e.g., social or cultural issues) are more likely to produce declarative, non-binding policies than IOs organized around issue areas where regulatory, binding output is central to the IO's core function (e.g., trade).

Identifying such patterns of policy output across the five dimensions opens up the possibility of distinguishing different categories of IO performance. For instance, there may exist a category of "talking shop IOs," producing high volumes of declarative, non-binding policies with low target specificity, covering a diverse list of issue areas. At the opposite end of the spectrum, there may be a group of IOs whose performance is characterized by low-volume production of regulatory, binding policies in one issue area, targeting one single type of actor.

Implications for Institutional Design and Effectiveness

By way of conclusion, we broaden the perspective to address how knowledge about the performance of IOs may help to inform the study of institutional design and effectiveness in global governance. Once we have measured and mapped the policy output of IOs, we can begin to address some of the most critical issues at stake in the theory and practice of international cooperation.

First, information about the performance of IOs will make it possible to systematically explore the effects of *institutional design* on how organizations work. While existing research has explored the sources and patterns of international institutional design, there is so far limited systematic knowledge about the consequences of alternative institutional designs. Still, we expect there to be such effects, based on the widely held assumption that institutions

matter for actor behavior, whether by providing incentives or standards of appropriate behavior (Shepsle 1979; March and Olsen 1989). Data on IO performance would allow us to move from assumptions and hypotheses to empirical assessments. In that context, it would also be possible to evaluate alternatives to institutional design as sources of IO performance, notably, organizational culture, power distribution, and heterogeneity of state preferences (Gutner and Thomson 2010).

Based on existing research, several dimensions of the institutional design of IOs may be identified with a potential effect on their performance (cf. Koremenos et al. 2001; Hawkins et al 2006; Tallberg et al. 2013; Hooghe and Marks 2014; Knill and Bauer 2015). Transnational access, in terms of involvement of non-state actors, may either improve IO performance by contributing resources, or reduce efficiency by hindering effective decision-making. Autonomy for international bureaucracies, in terms of independence from member-state control, may either make IOs more able to advance policy, or more prone to dysfunctional organizational behavior. Policy scope, in terms of the breadth of mandates, may either make IOs perform better, or make them more unfocused. Agenda-setting rules, by privileging one or multiple agenda setters, may impact the stability, efficiency, and richness of policy adoption. Decision rules, in terms of thresholds for policy adoption, may affect the efficiency of decision-making, or lack importance because of a primacy of informal norms.

Second, policy output likely matters for the *effectiveness* of IOs (Underdal 2002; Gutner and Thomson 2010; Young 2011). Unless IOs are capable of developing and adopting policy aimed at solving the transnational problems societies confront, they will be of limited value in combatting environmental degradation, intrastate conflict, financial instability, human-rights abuse, and other challenges. As we put it earlier, policy output is a first step and a necessary condition for IOs to have an effect. At the same time, it is clear that IO policy adoption in no way guarantees an impact on the problems addressed, since policies may be insufficiently

ambitious, badly designed, undermined by non-compliance, or counteracted by contextual factors beyond political control. Hence, the precise relationships between policy output and problem-solving effectiveness are complex and remain to be disentangled. Careful conceptualization, measurement, and mapping of policy output will be vital to that project.

An important conclusion in existing scholarship is that IO performance tends to be a more important determinant of problem-solving effectiveness than measures of whether the problem is benign (easy to solve) or malign (difficult to solve) (Miles et al. 2002; Young 2011). Even malign problems can be addressed effectively if IOs are capable of developing and adopting well-designed policies. Another central conclusion is that policy designs which work well in one context often do not work equally well elsewhere, even if we stay within the same issue area. One-size-fits-all solutions are rare and sensitivities to contextual factors high. Taken together, these conclusions underline the need to conceptualize IO performance in terms of multiple dimensions, which may interact in producing more or less effective policy solutions. Our typology of policy output in terms of five main features is useful in this regard.

References

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

Alesina, A., Angeloni, I. and Schuknecht, L. (2005) 'What does the European Union do?', *Public Choice* 123 (3-4): 275–319.

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.

Alvarez, R.M., Garrett, G. and Lange, P. (1991) 'Government partisanship, labor organization, and macroeconomic performance', *American Political Science Review* 85(2): 539–56.

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.

Bratton, K.A. and Haynie, K.L. (1999) 'Agenda setting and legislative success in state legislatures: The effects of gender and race', *Journal of Politics* 61(3): 658–79.

Carrubba, C. and Gabel, M. (2014) *International Courts and the Performance of International Agreements: A General Theory with Evidence from the European Union,* Cambridge: Cambridge University Press.

Cichowski, R.A. (2006) 'Courts, rights, and democratic participation', *Comparative Political Studies* 39(1): 50–75.

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.

Easton, D. (1964) *A Framework for Political Analysis*, Englewood Cliffs: Prentice Hall.

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

Fioretos, O. (2011) 'Historical institutionalism in international relations', *International Organization* 65(2): 367–99.

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.

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.

Hafner-Burton, E M. (2013) *Making Human Rights a Reality*. Princeton: Princeton 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. (2014) Delegation and Pooling in International Organizations. Online First, *Review of International Organizations*.

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.

Keohane, R.O. and Victor, D.G. (2011) 'The regime complex for climate change', *Perspectives on politics* 9(1): 7–23.

Knill, C. and Bauer, M. W. (2015) 'Introduction: Looking into the black box of global governance: the role of international bureaucracies in policy-making. Unpublished paper'.

Koremenos, B., Lipson, C. and Snidal, D. (2001) 'The rational design of international institutions', *International Organization* 55(4): 761–99.

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

Lipson, C. (1991) 'Why are some international agreements informal?', *International Organization* 45(4): 495–538.

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.

Mansfield, E.D. and Reinhardt, E. (2008) International institutions and the volatility of international trade', *International Organization* 62(4): 621–52.

March, J.G. and Olsen, J.P. (1989) *Rediscovering Institutions: The Organizational Basis of Politics*, New York: Free Press.

Marin-Bosch, M. (1987) 'How nations vote in the General Assembly of the United Nations', *International Organization* 41(4): 705–24.

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.

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. (1994) 'Creeping competence: the expanding agenda of the European Community', *Journal of Public Policy* 14(2): 95–145.

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. (2012) *International Organization*, 2nd edition, Basingstoke: Palgrave Macmillan.

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

Scharpf, F.W. (1988) 'The joint□decision trap: lessons from German federalism and European integration', *Public Administration* 66(3): 239–78.

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

Scruggs, L.A. (1999) 'Institutions and environmental performance in seventeen western democracies', *British Journal of Political Science* 29(1): 1–31.

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.

Shepsle, K. A. (1979) 'Institutional arrangements and equilibrium in multidimensional voting models', *American Journal of Political Science* 23: 27-60.

Swedish Institute of European Policy Studies (SIEPS) (2015) 'Statistical Survey of the Legislation in the EU', Available online at: http://sieps.se/en/forskningsprojekt/statistical-survey-of-the-legislation-in-the-eu.

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

Thomson, R., Stokman, F.N., Achen, C.H. and König, T. (eds.) (2006) *The European Union decides*, 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.

Underdal, A. (2004) 'Methodological challenges in the study of regime effectiveness. In Regime Consequences', in A. Underdal and O.R. Young (eds.) *Regime Consequences:*

Methodological Challenges and Research Strategies, Dordrecht: Kluwer Academic Publishers, pp. 27-48.

Wettestad, J. (2002) 'The Vienna Convention and Montreal Protocol on ozone-layer depletion', *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. 149–70.

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

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).