

Coen, David; Kreienkamp, Julia; Tokhi, Alexandros; Pegram, Tom

**Article — Published Version**

## Making global public policy work: A survey of international organization effectiveness

Global Policy

**Provided in Cooperation with:**

WZB Berlin Social Science Center

*Suggested Citation:* Coen, David; Kreienkamp, Julia; Tokhi, Alexandros; Pegram, Tom (2022) : Making global public policy work: A survey of international organization effectiveness, Global Policy, ISSN 1758-5899, Wiley, Oxford, Vol. 13, Iss. 5, pp. 656-668, <https://doi.org/10.1111/1758-5899.13125>

This Version is available at:

<https://hdl.handle.net/10419/265248>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<http://creativecommons.org/licenses/by-nc/4.0/>

## Appendix I

Table A provides an overview of the 13 IOs included in our expert survey. Organizations marked with an asterisk have been selected by respondents for further in-depth evaluation in the second part of the survey. In the case of trade, experts selected the WTO as only relevant global IO (noting that UNCTAD is covered under investment in our survey).

*Table A List of surveyed IOs.*

<b>IO Name</b>	<b>Abbreviation</b>	<b>Issue Area</b>
UN Framework Convention on Climate Change*	UNFCCC	Climate
Intergovernmental Panel on Climate Change	IPCC	Climate
UN Development Program	UNDP	Development
World Bank*	WB	Development
Bank for International Settlements*	BIS	Finance
Basel Committee on Banking Supervision	BCBS	Finance
International Centre for Settlement of Investment Disputes*	ICSID	Investment
UN Conference on Trade and Development	UNCTAD	Investment
UN High Commissioner for Refugees*	UNHCR	Migration
International Organization for Migration	IOM	Migration
UN (Security Council) *	UNSC	Security
North Atlantic Treaty Organization	NATO	Security
World Trade Organization*	WTO	Trade

## Appendix II

### Questionnaire: IO Effectiveness Across Policy Areas

Before you begin the survey, please specify your issue area:

- ☐ Climate change
- ☐ Development
- ☐ Finance
- ☐ Investment
- ☐ Migration
- ☐ Security
- ☐ Trade

1. Please identify two key formal or informal international organizations (IOs) in your issue area and list their main policy goals (indicating source statute):

	Name of IO	Key policy objective(s)
1.		
2.		

#### Conceptualizing IO Effectiveness:

There are various approaches to defining and empirically assessing effective IO performance. A central question for a growing number of effectiveness scholars is whether IOs **actually change state (or non-state) behaviour in a way that contributes to the achievement of their policy goals**. Put simply, do international organizations actually make a difference? This necessarily involves some counterfactual reasoning: Are we seeing a change in behaviour that contributes to the achievement of stated policy goals and that would likely not have occurred in the absence of this IO? While this is clearly an important question, the process through which IOs influence the behaviour of rule-takers, directly or indirectly, is often long, complex and difficult to assess.

Therefore, others have considered policy outcome as a proxy for effectiveness. This conceptualisation of effectiveness focuses on **the degree to which an IO is able to influence the global agenda, bring together different stakeholders, and generate consensus on norms, goals and rules**. The degree of ambition reflected in norms, goals and rules is also important to gauge effectiveness. Does institutionalisation through the IO move norms, goals and rules beyond

the lowest common denominator? While such output in itself can be considered a reflection of IO effectiveness (“constitutive effectiveness”), it does not deliver insights on whether states follow up on their commitments.

Thus, another group of institutionalist scholars has framed IO effectiveness primarily in terms of their **ability to monitor and enforce compliance with international rules**. The work of these scholars has helped illuminate the institutional effects of IOs, yet compliance alone does not fully capture whether multilateral institutions actually induce a change in state behaviour that would not have otherwise occurred.

Because all three conceptualisations of IO effectiveness (goal attainment, constitutive effectiveness, and compliance) raise difficult theoretical and methodological challenges, in this inquiry, we consider them in combination.

For more information, please refer to the enclosed executive summary.

2. **Goal attainment:** How well are the two IOs identified above doing in terms of achieving their stated policy goals? Please expand below, starting with the IO you consider most effective in terms of goal attainment.

	Name of IO	Assessed performance in terms of goal attainment
1.		
2.		

Additional comments (e.g. on how goal attainment differs across multilateral agreements concluded under the mandate of the IOs above):

3. **Constitutive effectiveness:** How well are these two IOs doing in terms of generating consensus on (ambitious) shared norms, goals and rules? Please expand below, starting with the IO you consider most effective in terms of constitutive effectiveness.

	Name of IO	Assessed performance in terms of constitutive effectiveness
1.		

2.		

Additional comments (e.g. on the tension between ambition and broad-based consensus):

4. **Compliance:** How well are these two IOs doing in terms of monitoring and enforcing rule compliance? Please expand below, starting with the IO you consider most effective in terms of inducing compliance.

	Name of IO	Assessed performance in terms of compliance
1.		
2.		

Additional comments (e.g. on how compliance outcomes differ across multilateral agreements concluded under the mandate of the IOs above):

### Conceptualizing IO authority:

IO authority has two dimensions: autonomy and bindingness. Thus, it reflects both the degree of independence an IO enjoys in decision making and the extent to which its decisions can limit state discretion.

IOs may exercise authority across different policy functions, including (but not necessarily limited to):

- **Agenda setting:** determining which policy proposals are included or excluded for debate and deliberation
- **Rule making:** translating policy proposals into substantive obligations or guidance for state parties
- **Compliance monitoring:** collecting and assessing information on state parties' compliance and performance
- **Norm interpretation:** passing judgments in case of disagreement over specific institutional norms or efforts to solve intra-state disputes (dispute settlement)
- **Enforcement:** imposing sanctions on non-compliant state parties
- **Evaluation:** assessing the IO's own performance and developing proposals to improve its internal workings
- **Knowledge generation:** collecting, processing, evaluating, and disseminating knowledge pertaining to the substantive governance problem the IO is tasked to address

How much authority an IO exercises across these policy functions is dependent on specific **institutional design features**, i.e. the processes, mechanisms and structures that allow the IO to perform authority across the policy functions listed above. These may include (but are not limited to):

- Mechanisms to ensure sufficient funding
- Safeguards to ensure independence
- Mandate prerogatives
- Regular reporting requirements
- Formal powers of investigation
- Powers to impose legal sanctions
- Dispute settlement and adjudication mechanisms
- Possibilities of engaging transnational actors in monitoring and other policy functions ("orchestration")

More general design features, such as size and structure of an IO, may also affect how much authority it enjoys across functions.

More information is contained in the enclosed executive summary. For a more extensive discussion, please refer to Zürn, M., Tokhi, A. and Binder, M. (2021). The International Authority Database. *Global Policy*. DOI: 10.1111/1758-5899.12971.

5. Pick the IO which you consider most authoritative in your issue area. For this IO, rank the following **policy functions** in order of their importance for IO effectiveness, from 1 (least important) to 7 (most important). See above for further information on policy functions.

Name of IO:

Ranking of policy functions:

- Agenda setting Choose an item.
- Rule making Choose an item.
- Compliance monitoring Choose an item.
- Evaluation Choose an item.
- Enforcement (sanctioning) Choose an item.
- Norm interpretation (dispute settlement) Choose an item.
- Knowledge generation Choose an item.

Is there another policy function that is important to effectiveness of this particular IO? If so, what is it and how highly would you rank it in terms of importance?

Additional comments:

6. For the same IO, which **design features** are particularly conducive to goal attainment? Please list up to five design features in order of importance. These may include design features listed in the box above or others that you consider important.

	Design feature	Further explanation
1.		
2.		
3.		
4.		
5.		

Additional comments:

7. Again, for the same IO, which existing **policy function(s)** would need further improvement to enhance effectiveness in your view? Please choose from the below (1 = policy function needs enhancement, 0 = policy function is sufficient as it is).

- Agenda setting Choose an item.
- Rule making Choose an item.
- Compliance monitoring Choose an item.
- Evaluation Choose an item.
- Enforcement (sanctioning) Choose an item.
- Norm interpretation (dispute settlement) Choose an item.
- Knowledge generation Choose an item.

If you could pick just one **policy function** that most requires improvement to enhance effectiveness of the IO, which one would it be and why? Which particular **design features** could help improve this function?

8. In your opinion, what are the most important overall challenges to effectiveness of IOs in your issue area?

9. Please provide a general abstract (up to 500 words) laying out the current state of IO effectiveness in your issue area.



## Appendix III

To arrive at a more systematic comparison of IO effectiveness, we coded experts' responses to the open-ended questions (2, 3, and 4) in the questionnaire. Specifically, we aggregated responses into an ordered variable, expressing whether IO effectiveness is considered high, moderate, or low. Table B below shows the distribution (in percentages) of ordinal ratings for each dimension of effectiveness and in each issue area.

In what follows, we briefly outline our coding of expert assessments on IO effectiveness. We defined three ordered values that describe the degree of IO effectiveness. Effectiveness is low when experts evaluate the IO as barely capable fulfilling its mandate, highlighting important deficiencies in the organization's set-up and operation. Effectiveness is moderate when the IO is effective in some areas and tasks, but shows notable difficulties in achieving progress on other facets of its mandate. Finally, we code IOs as highly effective when experts explicitly mention that the organization is overall effective and does not witness major obstacles in pursuing its mandate. Here, we discuss the specific coding rules, along with illustrative examples, for each of the three effectiveness dimensions.

*Constitutive effectiveness.* To code IOs' constitutive effectiveness, we focused on two sub-components, namely (1) the degree to which an IO achieves normative consensus on ambitious policy goals among different stakeholders and (2) the extent to which it influences global policy agendas. Experts' qualitative assessments allowed us to extract the relevant information and rate the performance of each organization with respect to these sub-components. For example, the UNFCCC has been evaluated as successful in achieving normative consensus and thus we assign the value high on sub-component 1. By contrast, with respect to influencing a shared global policy agenda, the UNFCCC has only partially managed to do so. Accordingly, we rate it as moderately effective on sub-component 2. The resulting final value on constitutive effectiveness is moderate, given that the IOs achieved only partial progress in this area. If an IO scores high on both sub-components, its resulting final constitutive effectiveness score is high. By contrast, if both sub-components receive the value low, its final score is low.

*Compliance.* We proceeded similarly with respect to coding compliance. Experts identified at least

one policy/obligation per IO and rated its ability to induce state compliance. Qualitative responses allowed us to identify whether IOs achieve high, moderate, or low levels of compliance regarding the indicated policies/obligations. For example, experts identified significant gaps and systematic non-compliance with NATO's goal of raising members' defense expenditures to 2 percent of their GDP. Accordingly, compliance with that policy is low. By contrast, most states follow the IOM's border control regime and the organization is effective in implementing the associated rules. Accordingly, compliance is rated as high in this case. The UNHCR has been partly successful in promoting compliance with the Geneva Conventions of 1951/1967 through facilitative mechanisms, but lacks instruments to monitor state behavior with the result that some of its provisions are ignored by states. Accordingly, we rate the UNHCR as moderately effective in eliciting compliance.

*Goal attainment.* Our coding of IOs' goal attainment is analogous to the other dimensions. Experts evaluated at least two policy goals per IO, each of which we coded using our ordinal rating scale. For example, experts highlighted that the UNSC achieves partial progress in protecting international peace and security through its peacekeeping missions (policy goal 1) and, likewise, has only moderately contributed to halting the spread of weapons of mass destruction through UNSC Resolution 1540 (policy goal 2). Given this mixed record, the UNSC is assigned the final value moderate in this dimension of effectiveness. By contrast, the BIS has been judged to have effectively coordinated an informal network of central banks (policy goal 1) and advanced the harmonization of their financial policies (policy goal 2). Accordingly, we code the BIS's goal attainment as highly effective.

Two members from our group independently coded the expert assessments. In cases of conflict (3 out of 35 ratings), we discussed the conflict and consulted the overall assessment of IO effectiveness provided in question 9 of the questionnaire to arrive at a final score. Our coding scheme is not an exact mapping of qualitative accounts to numbers, but rather serves to systematize and compare the rich expert accounts on IO effectiveness.

In total, we arrive at 35 distinct ratings of IOs. More specifically, in the case of constitutive effectiveness and goal attainment, 13 IOs are rated in each dimension ( $13 \times 2 = 26$  ratings). However, four out of 13 IOs have no compliance rating, because the measure of compliance does not apply in their case. Accordingly, only 9 IOs were coded with respect to their ability to induce state compliance,

resulting in a total of 35 IO ratings ( $26 + 9 = 35$ ).

Table B summarizes expert ratings. We group IOs from the seven issue areas into five distinct clusters. Cross-tabulating the issue clusters with the effectiveness dimension, we further differentiate between the three levels of our rating variable. Cell entries show percentages of rated IOs within each combination of rows and columns. For example, all climate IOs (100%) have been rated as moderately effective with regard to their constitutive effectiveness. By contrast, none of the Finance & Investment IOs has been rated as low-effective with regard to their goal attainment. Row totals indicate for the IOs in each issue area the distribution of rating scores. For example, in Trade & Development 12.5 of IOs have been rated as moderately effective.

*Table B Percentage of coded expert ratings across dimensions of effectiveness and issue areas.*

	<b>Rating</b>	<b>Constitutive</b>	<b>Compliance</b>	<b>Goal Attainment</b>	<b>Total</b>
<b>Climate</b>	High	0	0	50	20
	Moderate	100	0	50	60
	Low	0	100	0	20
<b>Trade &amp; Development</b>	High	100	100	66.6	87.5
	Moderate	0	0	33.3	12.5
	Low	0	0	0	0
<b>Finance &amp; Investment</b>	High	50	0	25	30
	Moderate	25	100	75	60
	Low	25	0	0	10
<b>Security &amp; Migration</b>	High	25	25	0	16.6
	Moderate	25	50	100	58.3
	Low	50	25	0	25.1
<b>Total</b>	High	46.1	33.3	30.7	37.1
	Moderate	30.7	44.4	69.2	48.6
	Low	23.2	22.2	0	14.3

## Appendix IV

In this section, we investigate how specific design features (processes, mechanisms and structures) highlighted by survey respondents boost IOs' authority within or across a range of policy functions, thus contributing to goal attainment. This is based on question 6 in our survey, which asked experts to identify and briefly describe up to five design features that are particularly conducive to the overall performance of the IO that they consider most authoritative in their issue area. Experts could peruse a list of examples and also make their own suggestions, allowing for a degree of flexibility.

To help us systematize answers, Table B maps these features onto the seven policy functions introduced earlier in the survey (agenda setting, rule making, compliance monitoring, norm interpretation, enforcement, evaluation, knowledge generation), where applicable. For example, regular reporting requirements or the power to authorize on-the-ground inspections are institutional mechanisms contributing to compliance monitoring. We coded answers only slightly to allow for better comparison, specifically where mentioned design features contributed to more than one policy function. Notably, experts frequently mentioned design features, such as convening powers or provisions to safeguard independence, which enhance an IO's role as an active governance leader in its issue area (coded 'focality') and may thus simultaneously support a range of policy functions. For example, in the case of the World Bank, focality refers to its ability to act relatively autonomously from states when designing and implementing new projects, leveraging its stock of past experiences and technical expertise. Thereby, focality supports three policy functions, namely agenda setting (discussing and determining project priorities), rule making (adopting projects), and knowledge generation (documenting its experience, producing new insights, and using them for its project design). Similarly, design features that allow IOs to coordinate activities by intermediary actors (coded 'orchestration') or safeguard financial resources (coded 'funding') may also contribute to several or, in the latter case, all policy functions.

Table C IO-specific design features mentioned in the surveys and corresponding policy functions considered conducive to IO goal attainment

	Mentioned design features	Corresponding Policy functions
<b>Climate: UNFCCC</b>	Regular reporting requirements	Compliance monitoring
	Orchestration of sub-national and non-state actors ( <i>orchestration</i> )	Several (incl. enforcement / compliance monitoring)
	Competence to convene COPs, ambition ratcheting ( <i>focality</i> )	Several (incl. agenda setting / rule making)
<b>Development: World Bank</b>	Mechanisms to ensure sufficient funding ( <i>funding</i> )	All
	Independence to design new projects based on past experience and learning ( <i>focality</i> )	Several (incl. agenda-setting / rule-making / knowledge generation)
	Regular reporting requirements	Compliance monitoring
	Powers to impose legal sanctions	Enforcement
<b>Finance: BIS</b>	Standard-setting powers ( <i>focality</i> )	Several (incl. agenda setting / rule-making / knowledge generation)
	Mechanisms for norm interpretation	Norm interpretation
<b>Investment: ICSID</b>	Dispute settlement and adjudication mechanisms	Norm interpretation
<b>Migration: UNHCR</b>	Mechanisms to ensure funding ( <i>funding</i> )	All
	Safeguards to ensure independence and strengthen Executive Commission ( <i>focality</i> )	Several (incl. agenda setting / rule-making)
	Coordinate different stakeholders ( <i>orchestration</i> )	Several (incl. enforcement / compliance monitoring)
<b>Security: UN</b>	Mandate prerogatives ( <i>focality</i> )	Several (incl. agenda setting / rule-making)
	Powers to impose legal sanctions	Enforcement
	Mechanisms to ensure sufficient funding ( <i>funding</i> )	All
	Regular reporting requirements / formal powers of investigation	Compliance monitoring
<b>Trade: WTO</b>	Dispute settlement and adjudication mechanisms	Norm interpretation
	Powers to impose legal sanctions	Enforcement
	Decision-making structures ( <i>focality</i> )	Several (incl. agenda setting / rule-making)

## Appendix V

In this section, we briefly present the main challenges to overall effectiveness reported per issue area and the codes we assigned to each challenge in order to systematize responses. We inductively extracted seven codes based on expert responses. These are: Ambition vs. broad-based consensus, Problem structure, Institutional capacity and finance, Fragmentation and lack of central authority, Political conflict, Legitimacy and representation, and Internal culture. Table C maps reported challenges onto these codes. For example, a lack of inclusion of non-western experts in global finance governance points to broader issues of representation and legitimacy.

*Table D Reported challenges to IO effectiveness and assigned codes.*

Issue Area	Reported challenge	Code
<b>Climate</b>	Goal ambition vs breadth of participation	Ambition – Broad-based consensus
	Effectiveness vs equity/justice	Legitimacy and representation
	Issue complexity	Problem structure
	Distributive consequences of climate action	Political conflict
<b>Development</b>	Competing development models	Problem structure
	Tension economic vs sustainable development	Problem structure
	Outsized influence of developed member states	Legitimacy and representation
	Inadequate indicators of development	Problem structure
	Local antagonisms re development projects	Political conflict
	Conflicting policy priorities	Political conflict
<b>Finance</b>	Inclusion of non-western experts	Legitimacy and representation
	Inclusion of social and consumer interests	Legitimacy and representation
	Risk analysis on social impact and climate change issues	Problem structure
	Regime fragmentation	Fragmentation and lack of central authority
	Lack of formal authority and legitimacy	Fragmentation and lack of central authority
<b>Investment</b>	Fragmentation and lack of authoritative IO to form central core of the regime	Fragmentation and lack of central authority
<b>Migration</b>	A generally hostile and increasingly restrictive international environment vis-à-vis migration, particularly asylum/refugee protection agendas	Political conflict
	Insufficient mechanisms for monitoring and enforcing compliance with international laws and norms	Institutional capacity and finance

	UNHCR internal culture, which is characterized by a lack of learning, deference to seniors, and deference to states	Internal culture
<b>Security</b>	Political disagreements and divergent national agendas	Political conflict
	Fast-changing nature of security threats and military material	Problem structure
	Operational constraints such as interoperability and funding challenges	Institutional capacity and finance
	Representation in UNSC	Legitimacy and representation
<b>Trade</b>	High politicization and contentiousness of trade policy	Political conflict
	Tension between ambition and consensus in rule-making	Ambition vs broad-based consensus
	High economic stakes of trade rules	Political conflict
	Global questioning of free trade (protectionism)	Legitimacy and representation

## Appendix VI

Table D presents the ranking of policy functions as calculated from data in the International Authority Database (IAD) (Zürn, Tokhi, Binder, 2021). For each function, we calculate its corresponding authority score. That score ranges between zero (no authority) and one (full authority). The IAD sample consists of 34 IOs, representative of the geographic and thematic distribution of IOs. The average authority score per function has been calculated for the year 2013. The average authority level across all functions is 0.28 and the standard deviation 0.35. The ranking in Table D broadly aligns with the relative importance of policy functions assigned by experts in our survey. The only exception is compliance monitoring.

*Table E Ranking of policy functions according to the data in the IAD.*

<b>Policy function</b>	<b>Rank</b>	<b>Authority Score</b>
Agenda setting	1	0.50
Norm interpretation	2	0.42
Rule making	3	0.34
Knowledge generation	4	0.28
Compliance Monitoring	5	0.20
Enforcement	6	0.17
Evaluation	7	0.08