Introduction: The Power of Global Performance Indicators

Judith G. Kelley
Dean and Terry Sanford Professor
Duke Sanford School of Public Policy
Judith.Kelley@duke.edu

Beth A. Simmons
Andrea Mitchell University Professor of Law, Political Science and Business Ethics
University of Pennsylvania
simmons3@law.upenn.edu

Abstract:

In recent decades, IGOs, NGOs, private firms and even states have begun to regularly package and distribute information on the relative performance of states. From the World Bank's Ease of Doing Business Index to the Financial Action Task Force blacklist, Global Performance Indicators (GPIs) are increasingly deployed to influence governance globally. We argue that GPIs derive influence from their ability to frame issues, extend the authority of the creator, and — most importantly — to invoke recurrent comparison that stimulate governments' concerns for their own and their country's reputation. Their public and ongoing ratings and rankings of states are particularly adept at capturing attention not only at elite policy levels but also among other domestic and transnational actors. GPIs thus raise new questions for research on politics and governance globally. What are the social and political effects of this form of information on discourse, policies and behavior? What types of actors can effectively wield GPIs and on what types of issues? In this symposium introduction, we define GPIs, describe their rise, and theorize and discuss these questions in light of the findings of the symposium contributions.

Introduction: The Power of Global Performance Indicators

Judith G. Kelley and Beth A. Simmons

Political leaders care about their reputations. Moreover, they appear to care about their reputations relative to their peers. Entrepreneurial actors have seized on this as an opportunity for policy influence. They have taken advantage of the competitive pressures of globalization, the increasing accountability demands flowing from democratization, and the transformative capacities of new information technologies to nudge global governance in important ways by deploying what we call *Global Performance Indicators* (GPIs) – regularized public assessments that rate, rank and categorize state policies, qualities, and/or performance. Such GPIs aim to (re)define global standards, contrive competition and provoke action according to specific performance criteria. They exemplify how states, intergovernmental organizations (IGOs), nongovernmental organizations (NGOs), think tanks and other states seek to use modern information politics to invoke reputational concerns and influence policy *through the pressure of comparison*.

Performance assessments and associated indicators permeate human activity, internationally and nationally, across public and private spheres. Some, such as product ratings in *Consumer Reports*, or national credit ratings by Standard and Poor's (S&P), have been around for decades; many are intended to inform private decision making. More recently, agenda-motivated indicators such as the country ratings of *Freedom in the World* by Freedom House or the *Ease of Doing Business* Index by the World Bank have appeared. Such indicators overtly and purposively seek to engage actors at the highest levels of government and influence their policies and governance practices.

The rising use of indicators has prompted debates about their construct validity, use and misuse.[1] Scholarship has flourished about methodology, reliability, and fairness for a broad range of indicators and have problematized quantification generally.[2] We build on a range of insights – from sociology to anthropology to political science – but the symposium focuses on intentional efforts to use comparative information to influence policy and governance. Why do states and other important actors care about ratings and rankings? Is effective use of GPIs confined to already-powerful actors such as the US, or can GPIs contribute to a diffusion of authority? And if a GPI is effective, does it work primarily through material leverage, or are other mechanisms, such as reputational or social pressure, at work as well?

This symposium expands our ability to answer these questions by examining the effects of four GPIs: The World Bank's *Ease of Doing Business (EDB) Index*, the *Millennium Development Goals* (MDG) for education, the Financial Action Task Force's (FATF) "blacklist" of non-complying jurisdictions, and an NGO called Publish What You Fund's *Aid Transparency (AT) Index*. The papers present evidence that these GPIs influence the discourse, standards, and measurement, but also examine whether and how they change actions. Do they reverberate in domestic politics, provoke bureaucracies, shape policy agendas and influence outputs? Do they influence the reactions and responses of donors and market actors? If they alter priorities, do they have *unintended* consequences as well? This symposium aims to speak directly to how GPIs might alter politics, policies and, more speculatively, even broader authority structures.

---

[1] Malito, Umbach, and Bhuta 2018.

[2] Arndt et al. 2006 ; Ginsburg et al. 2005 ; Høyland, Moene, and Willumsen 2012 ; van de Walle 2006. The World Bank's *Governance Indicators* are a potent example. See Apaza 2009. For a critique of their construct validity see Thomas 2010.

I.   Global Performance Indicators: definition, scope, and context

Global performance indicators are defined as *a named collection of rank-ordered data that purports to represent the past or projected performance of different units*.[3] While thousands of basic indicators such as Gross Domestic Product (GDP) exist to facilitate research or policymaking, and while all such quantification is normative and contestable,[4] this symposium focuses on the consequences of overtly strategic state rating and ranking systems that package and deploy information intentionally for policy advocacy and implementation. Such GPIs assemble selected data – facts and statistics collected together for reference or analysis – to craft "new" information designed and publicized to complement advocacy for specific ends. We focus on GPIs that, in their purest form, meet all the following criteria:

- **Public** and easily available.

- **Regular** and published on a predictable schedule.

- **Purposive**, explicitly normative, policy focused

- Deployed to influence **state-level outcomes**.

- **Comparative** of the performance of multiple states within a region or more broadly.

GPIs can take several forms. ***Indexes or indicators*** use numbers or grades to rate or rank state performance, compressing enormous variance into a simplified scale.[5] ***Categorical assessments*** use ordinal categories to produce (un)flattering peer groups, while ***Blacklists and Watchlists*** draw stark distinctions between compliers and offenders. Hybrid systems are also common, as are accompanying country narratives and policy recommendations.

---

[3] Davis et al. 2012 ; Davis, Kingsbury, and Merry 2012 ; Davis, Merry, and Kingsbury 2015 ; Merry and Conley 2011.

[4] Fioramonti 2013.

[5] Broome and Quirk 2015 ; Büthe 2012 ; Davis et al. 2012 ; Hansen 2012 ; Karpik 2010.

Basic indicators are hardly new – sovereign credit ratings first appeared in 1916 – but nearly all the GPIs that fit our definition were created after 1990.[6] To demonstrate the trend, we systematically gathered a database based on secondary literature, extensive searches of newspapers and magazines and the wider web, using terms such as ranking, rating, index, blacklist, and watchlist. Each entry was coded among other things on launch year, regularity, purpose, creator type, country of origin, issue area(s), and the type of comparison used. To be included, a GPI had to have been created no later than 2015. This search produced 282 entries. Extensive culling[7] left 159 entries that met our criteria, and of these, 133 were still active as of 2018, meaning they had had been updated in 2015 or later.[8] These 159 were included in the source data for Figure 1. The dataset is available in the online appendix.

GPI growth has been nearly exponential. While it took almost three decades to approach 20 GPIs in the late 1990s, by the next decade the number had roughly quadrupled, and then in the next 15 it more than tripled. This proliferation responded to growing demand for policy-relevant performance data, spurred by a consensus that many past development reforms and governance structures were failing,[9] and facilitated by the fact that information was becoming ever easier to collect, process and disseminate.

---

[6] Löwenheim 2008.

[7] Of the 289, 29 were never reissued, so failed the regularity criteria. Another 28 were aimed at investors, so failed the policy purpose. Yet 30 had no explicit intent stated, so while they might qualify, they fell in a gray zone and for clarity of criteria, we noted them separately. Another 31 were just data (and clearly, we could have found hundreds more that would fit this category). A few others fell in other categories that didn't fit.

[8] The data collection concluded in July 2018.

[9] Arndt 2008 ; Arndt et al. 2006. Governance measures are the major theme in the Malito et al (2018) essay collection.
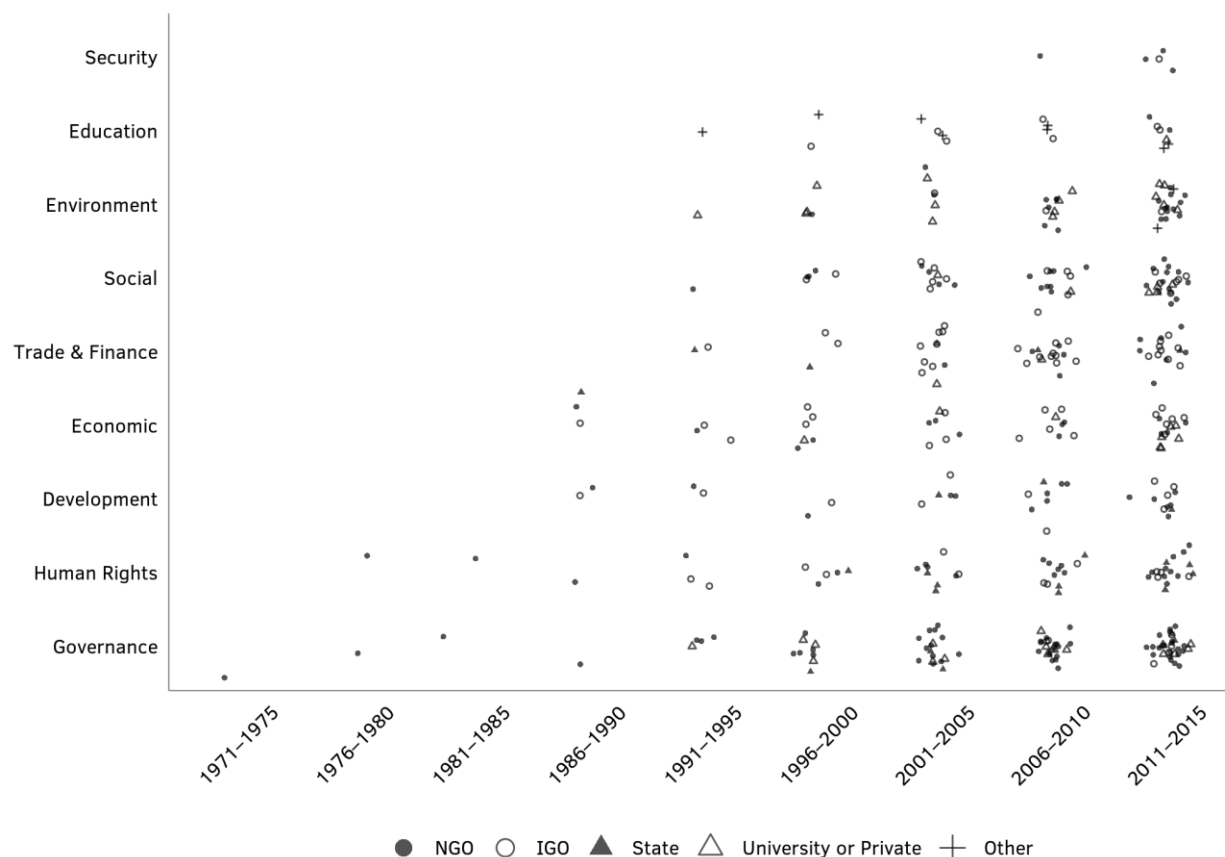
Figure 1: GPI Contestation across issues, time and creators. Source: Authors' database.

Notes: "Environment" covers Energy and environment. "Human rights" covers human rights, gender issues and press freedom. "Trade & Finance" covers Trade, Finance and Technology. Social covers health and social issues. All other topics are as coded in the dataset in the appendix.

As both the demand and supply of indicators grew, various actors began to realize that they could deploy GPIs as policy tools. A host of new GPIs, ranging from Transparency International's *Corruption Perceptions Index* to the OECD's *Program for International Student Assessment (PISA) rankings* began to flourish. Their labels became more normative and vivid; their content ever more actionable; their messages increasingly displayed to deliver visual impact. Recent GPIs are relentlessly comparative, suggesting an intention to pressure, shame or

provoke competition among states. Competitive prompts are clear in many GPIs' publicity messages. For example, when releasing the latest PISA rankings, the OECD announced: "Singapore tops latest OECD PISA global education survey."[10] Thus, GPIs are increasingly deployed to stimulate state competition and shape policy agendas; that is, as "technologies of power."[11]

II.   A theory of GPI power

Why should governments care about a simple ranking or rating, especially if it merely reformats preexisting information? We theorize that the power of GPIs is a function of their ability to engage reputations. Ratings and rankings stoke politicians' and bureaucrats' reputational concerns by framing, establishing "standards," and repeatedly engendering public comparisons. The packaging and intentional deployment of comparative information is thus an inherently social process that stimulate reputational concerns with how others view the performance of the state elites' competence, effectiveness and status. These concerns can operate through multiple pathways, depending on local values and institutions.

Sources of GPI power: Information, Contestation and the Power to Frame

The social process of GPI influence begins with their ability to frame issues. GPI promulgators seek to influence targets by designing an attention-grabbing form of political communication. By naming an issue, coining a vocabulary for describing it, and creating

---

[10] OECD website http://www.oecd.org/education/singapore-tops-latest-oecd-pisa-global-education-survey.htm, last accessed July 21, 2018.
[11] Hansen 2012.

categories for its assessment, promulgators hope to impact discourse and ultimately policy.[12]

This influence is akin to what Davis et al call "knowledge effects"[13] and what Pierre Bourdieu

terms "symbolic power" — or the power to name and categorize.[14] In some cases, the process

begins by collecting new data; in others, existing data is presented in ways that prioritize new

objectives. Such efforts shape how publics, organizations and policymakers think about "what

constitutes legitimate social practice."[15]

GPIs are both products of, and inputs into, active value contestation. Nowhere is this

clearer than in the struggle to define, measure and rate "development" over recent years.[16] The

United Nation's *Human Development Index* was created in 1990, adding lifespan, education, and

low fertility to traditional GDP per capita measures to reflect new concerns with the *quality* of

life. A decade later (2000), the UN new *Gender Equality Index* looked at *disparities* between

men's and women's quality of life as the central conceptual issue.[17] The Fraser Institute

contested both of these trends with its rival *Index of Human Progress* (2001) that restored weight

to high-income levels and reflected the Institute's interest in supporting competitive markets.[18] In

2006, the Sustainable Society Foundation, an NGO headquartered in the Netherlands, touted its

eponymously named index to contest the notion that unbounded growth is desirable.[19] In short,

creators of GPIs frame successful performance in preferred directions, often crowding the market

with GPIs that jostle for agenda-setting visibility.[20]

---

[12] Scheufele and Tewksbury 2007.
[13] Davis, Merry, and Kingsbury 2015.
[14] Bourdieu 1989.
[15] Hansen 2012.
[16] For the history of the Human Development Index, see Stanton 2007. For the World Bank's Governance Indicators see Kaufmann, Kraay, and Zoido August 1999.
[17] United Nations, Gender Inequality Index, http://hdr.undp.org/en/content/gender-inequality-index-gii
[18] See https://www.fraserinstitute.org/sites/default/files/MeasuringDevelopmentIHP.pdf.
[19] See http://www.ssfindex.com/.
[20] Miller 2005.

Evidence of contestation across multiple GPI sectors is suggested in Figure 1. It shows the gradual uptake of GPIs in different issues areas. A lack of consensus may be one motivation for creating new GPIs and may partly explain the variation in density across issues. For example, while there is likely a greater demand for information about economic and financial performance which has led to a greater supply by actors keen to dominate such visible issues, it is also likely that the lack of consensus around what constitute favorable policies and ideal outcomes has contributed to the proliferation. Indeed, although we excluded indicators created simply to inform investors, over a quarter of the GPIs meeting our definition (36) focused on economic issues with several alternatives measures of prosperity and wellbeing to contest the traditional GDP. Many GPIs also are concerned with various education or social issues (29%), governance (20%), development (20%), and environment (17%), with several clusters covering hybrids of these issues as well. Many such hybrids represent efforts to contest narrow measures and advocate for broader conceptualizations of policy priorities. This is one reason the classification of GPIs into issue areas results in several that fall in multiple categories. Many are meant precisely to pull traditional conceptions such as growth in unconventional direction, for example toward justice.

Social Knowledge and Authority

A key source of GPI power is the credibility and authority of its creator.[21] What makes some GPI creators more authoritative than others? The social psychology literature suggests that one source of legitimate authority is trust, which, in turn, develops out of a perception that an

---

[21] On private authority in international relations, see Hall and Biersteker 2002.

actor is fair, knowledgeable, and/or competent.[22] This may explain why many GPI creators attempt to bolster their credibility by transparently describing and defending their index's methodology. As importantly, actors who themselves are thought to exemplify specific norms are more likely to be viewed as authoritative in judging them,[23] and to have "normative power," which Manners defines as the ability to shape conceptions of what is "normal."[24]

GPI creators also gain authority based on their (assumed) competence and expertise,[25] which in turn may be inferred from their tangible power and wealth. The assumption that global competence in governance, economic policy and human rights for example resides in the western democracies, while contestable, has reinforced their "authority" to rate others in these areas. This may explain why, according to our data, organizations headquartered in the Global North create about 98 percent of the GPIs in existence today.

Social knowledge and authority is also supported by the tangible and intangible advantages of network centrality. Non-state actors with access to decision makers are better able to mine information and deploy it to assert authority.[26] States also have important network resources useful in developing and deploying GPIs. For example, the United States Department of State, which creates the *Trafficking in Persons Report* and grades country performance, uses its embassies and NGO contacts around the world to tap into many sources of information.

---

[22] Espeland and Sauder 2007 ; Rieh 2002 ; Wilson 1983. Well-respected raters have accumulative trust over decades, so people may act on this "expert" information, even when it erroneous; See Simonsohn 2011. On the importance of legitimacy for the effectiveness of international organizations see Buchanan and Keohane 2006 ; Zürn 2004.

[23] Sikkink 1993.

[24] Manners 2002.

[25] Monks and Ehrenberg 1999.

[26] Hall and Biersteker 2002.

Actors centrally located in a social and political network are better able to set agendas[27] and impact information flows,[28] which facilitates both data collection and GPI dissemination.[29]

Both states and non-state actors can accrue social authority, for different reasons. States have advantages that flow from their preexisting organization, broad resource base, and in some cases their political legitimacy. In contrast, non-state actors' authority rests on *independence* from powerful political actors. An interesting case is that of the *Aid Transparency Index*, explored by Honig and Weaver. They show that a small meagerly resourced NGO used its political independence and first mover advantage to establish the gold standard for funding transparency. Political independence is plausibly the strength of several GPIs in our database: about half are created by NGOs, and the rest by universities, private actors, or a combination of actors. The United States is the only country that has created GPIs that target other states, which is only about 6 percent of the total. Overall, GPIs are largely *not* controlled by individual states or IGOs, although those created by the major IGOs and the United States attract more attention and on average may be more influential.

Judgment and comparison: provoking status concerns and competition

GPIs are consequential because they engage the reputation of the persons or entities being assessed. A r*eputation* refers to a widespread belief that a person or an organization can reliably be characterized in a particular way.[30] *Status* and reputation are related, but status is explicitly comparative: it refers to the *relative* social or professional standing of someone or

---

[27] Carpenter 2011. For related "Social impact theory" see Latané 1981.
[28] Borgatti and Cross 2003.
[29] Stone 2002.
[30] On the reputational impact of comparisons see Kelley 2017. Chapter 2.

something in a formal or informal social hierarchy.[31] Both reputation and status are

quintessentially social constructs: they are granted or accorded only by a social community. GPIs

confer status, which cannot be credibly claimed or manufactured unilaterally by the target. What

makes this form of information deployment potentially powerful is its ability to broadly impact

social beliefs about successful states and appropriate policies. It is deployed precisely as a form

of social pressure on targets to conform – or suffer the reputational consequences, real or

imagined.

Exactly *why* people value their or their country's status and reputation is hotly debated.

States and their elites likely seek to cultivate reputations as a form of strategic capital[32] to

improve their payoffs.[33] Some care about their reputations for reasons having to do with their

social role or identity.[34] The massive literature on shaming, for example, is built on the premise

that officials want to avoid damaging their personal, professional or organizational reputations.[35]

Bureaucrats, politicians, and their publics[36] also care about their own or their country's *relative*

standing, which can be hard to establish in a thinly institutionalized international environment.[37]

Status be parlayed into material benefits and/or be an end in itself. [38] International relations

research accords with similar theories in economics,[39] psychology[40] and sociology[41] in accepting

---

[31] Parks and Masaki (2016) find that policy makers rate comparative assessments as more effective than single country assessments.

[32] van Ham 2001.

[33] Tingley and Walter 2011.

[34] Dafoe, Renshon, and Huth 2014.

[35] On social pressures; see Checkel 2001 ; Johnston 2001. On shaming by states or IGOs see Joachim, Reinalda, and Verbeek 2008 ; Lebovic and Voeten 2006.;By non-governmental actors, see Hendrix and Wong 2013 ; Murdie and Davis 2012 ; Risse 1999.

[36] Rosen 2005 ; Wohlforth 2009.

[37] Dafoe, Renshon, and Huth 2014.

[38] Rege 2008. Some scholarship even suggests that status can propel costly policy choices, even war. See Dafoe, Renshon, and Huth 2014.

[39] Harsanyi 1966.

[40] Blader and Chen 2012.

[41] Ridgeway 2013.

that people are as motivated by status as by power or material goods, with complex consequences for attitudes and behavior.

GPI creators seek to engage this concern with status and reputation by leveraging comparative information among peers.[42] Explicit comparisons create contexts in which judgments are formed and identities are established and reinforced.[43] They foster "commensuration," or "the comparison of different entities according to a common metric," as a way of making highly-simplified sense of the world.[44] Labels like "rule of law," "freedom," "sustainable development," "peacefulness," or "political risk" imply that very different entities are, for the purposes of one concept, *comparable*. Numbers facilitate standard-making through the simplest possible means – averaging. [45] The recurrent nature of GPIs incentivizes governments to look to the next iteration of the index and worry about the reputational consequences of their next rating or ranking. Moreover, the media is particularly fond of reporting relative rankings; in numerous interviews GPI creators frankly acknowledge that they created such indexes precisely to attract media attention.[46] About two-thirds of the active GPIs employ explicit top-to-bottom ranking systems, and over a third create clear normative categories or performance tiers, usually in addition to a ranking or rating. These features render GPIs a potent tool for producing social control through the pressure of comparative information.

---

[42] In one exceptionally ambitious effort to exercise social control through ranking, China reportedly has pilot programs to rate each and every citizen according to a form of "social credit." See reports at http://www.bbc.com/news/world-asia-china-34592186.

[43] Ashmore, Deaux, and McLaughlin-Volpe 2004.

[44] Espeland and Sauder 2007 ; Espeland and Stevens 1998. Sinclair 2008. Schueth 2011. Processes of commensuration inform global benchmarking; see Broome and Quirk 2015.

[45] Weisband 2000.

[46] This theme was evident in a series of 23 interviews conducted by the authors in Washington DC, August 12-14, 2014.

Pathways to responses: Domestic politics, bureaucratic responses, and transnational pressures

This symposium explores how comparative rankings influence policy choices through multiple mechanisms.[47] Figure 2 illustrates how GPIs reverberate through different levels of politics: domestic, elite and transnational, and underscores the cyclical nature of iterative assessments.
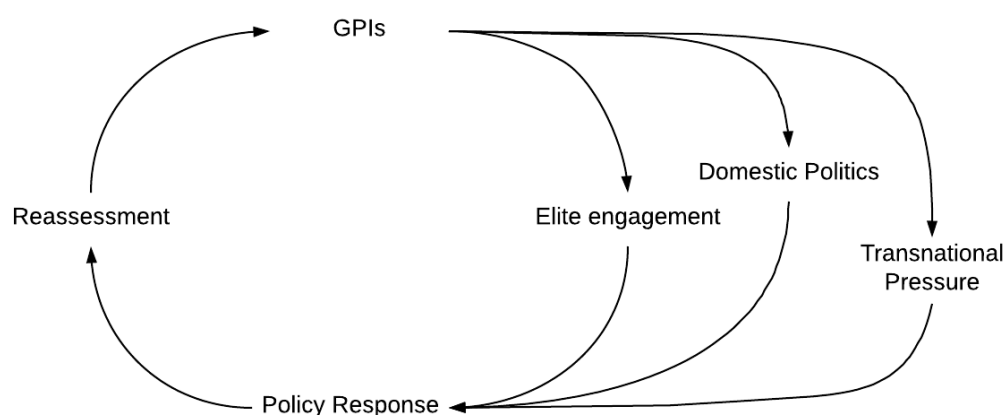
Figure 2: From GPIs to policy change: pathways, policy response and reassessment. *Adapted from Kelley and Simmons 2015 and Kelley 2017.*

First, GPIs can provide information that reverberates in domestic politics, especially when amplified by the popular or social media or circulated in civil society.[48] Important domestic audiences include local advocacy organization, local businesses, and even popular opinion generally. These groups use the comparative information from GPIs to demand better

---

[47] These processes align with seminal work on information politics, transnational activism and domestic mobilization including: Davenport, Eccles, and Prusak 1998 ; Heiss and Kelley 2017 ; Keck and Sikkink 1998 ; Ron, Ramos, and Rodgers 2005. In environmental policy, "tote-board diplomacy" has created pressures for change; see Levy 1993. We confirmed similar processes in a series of interviews with GPI creators in August 2014, Washington DC.

[48] Carpenter 2007 ; McCombs and Shaw 1972.

policy, especially when their country apparently lags behind. The underlying information need not be wholly new, but the *relative* performance is. The salience of the new information becomes an opportunity to mobilize in several ways: electorally through the ballot box, by publicizing critiques of performance in the media, engaging in traditional forms of lobbying, or by protesting or engaging in collective action. In responsive political systems, these demands may elicit policy change. At a minimum, officials will have the incentive to *claim* they are addressing the issue. However, where institutions repress public input and suppress political demands, governments may respond not with reform, but by denigrating the GPI or its creator.[49]

Articles in this symposium illustrate how GPIs resonate in domestic politics, creating new demands on governments. Honig and Weaver's research shows that the *Aid Transparency Index* enhances domestic political pressure by equipping transparency reform proponents (particularly those with some material power over aid agencies) with critical information and clear standards to guide policy change. Bisbee et al argue that the extra scrutiny associated with the MDGs criteria encourages greater domestic political attention to those criteria, creating demands for more investment in public goods that improve performance on the MDGs, even at the sacrifice of other equally important public investments. Businesses use GPIs to demand policy reform as well. Morse describes how Thai bankers and industrialists all urged the government to modify its laws and improve the image of the country's financial system in response to a non-compliance rating. Her research illustrates how a GPI can shape a domestic agenda, drawing attention to policy issues that might otherwise have been entirely disconnected from domestic preferences. In a twist, Doshi, Kelley and Simmons show how Modi used the *Ease of Doing Business* Index to provoke domestic awareness and pressure on his own

---

[49] See Kelley 2017, Chapter 4.

bureaucrats and provincial governments to move business reregulation along, effectively becoming an ally of the World Bank's agenda.

Second, many GPIs involve intensive consultations with government bureaucrats during which officials take "expert" advice directly into account in their policymaking.[50] Such monitoring signals the value the monitor and other actors attach to specific tasks,[51] and can influence targets to internalize monitoring regimes and eventually to *self*-regulate — a process that may mistakenly lead to underestimation of monitoring's effects.[52] Referred to as the "Hawthorne effect," individuals in experimental settings have been found to re-arrange their priorities to meet external expectations when they are aware of being observed.[53] This form of *reactivity*—the tendency for people to change their behavior in response to being evaluated— may explain the effect, for example, of US News and World Report's *Best Colleges* and *Best Grad Schools* rankings on university priorities.[54] When a monitoring regime is applied generally to like units, it eventually gains acceptance by undercutting claims that the monitors have singled out specific targets "unfairly."[55]

Used in conjunction with monitoring, GPIs directly engage the reputations for competence and effectiveness of specific officials and their agencies. When GPIs target policies for which specific ministries are directly responsible, officials initiate policy change to deflect criticism that could damage their reputations. As Honig and Weaver find, elites responded to the *Aid Transparency Index* by inter- and intra-organizational learning, adopting its norms, and professionalizing aid staff and management around the specific standards of transparency that the

---

[50] Parks and Masaki.(2016) find that GPIs on which bureaucrats are consulted are among the most salient to government officials. On the importance of bureaucrats' reputations, see Chwieroth 2013 ; Nelson 2017.

[51] Larson and Callahan 1990.

[52] Self-regulation has been found effective in the context of health and safety monitoring by OSHA; Johnson 2017.

[53] Adair 1984.

[54] Espeland and Sauder 2007.

[55] Löwenheim 2008.

index promotes. Doshi, Kelley and Simmons (this volume) note that some states have literally created agencies tasked to develop policies that meet the expectations of the World Bank, and show that politicians such as Modi use salient GPIs to justify the job they have done, take credit for progress, and set goals for the future.

Finally, GPIs activate transnational pressures and influence how third parties such as foreign investors, donors or other states respond to – or are *anticipated* to respond to – the ratings. Many familiar rating systems leverage social pressure to conform primarily through third parties: universities respond to the U.S News and World Reports rating system because applicants do;[56] hospitals respond because patients do;[57] firms respond to environmental rankings because investors do.[58] Importantly, raters themselves need not control material resources to unleash transnational pressures. Morse's research in this symposium shows that a non-cooperative designation by FATF on anti-money laundering policies can lead international banks to be less willing to lend money to banks in listed countries. Interestingly, however, Bisbee et al's examination of budget decisions in response to MDG goals shows that while primary education aid flows respond to the MDGs, there is no evidence that aid flows mediate the government's own response to the MDGs. Rather, the MDGs appear to work through domestic information channels.

III.   What have we learned?

---

[56] Luca and Smith 2013 ; Meredith 2004 ; Monks and Ehrenberg 1999.
[57] Pope 2009.
[58] Aaron, McMillan, and Cline 2012 ; Murguia and Lence 2014. Murguia and Lence found that getting one position closer to the top of *Newsweek*'s "Global 100 Green Rankings" increases the value of an average firm in the list by eleven million dollars.

The contributions in this symposium apply an eclectic set of methods to greatly expand our knowledge about how GPIs influence states' policies. Important decisions ride on such rankings, from significant changes in the criminal code (Morse) to increased aid transparency (Honig and Weaver) to allocation of school funding (Bisbee et al) to business deregulation (Doshi, Kelley and Simmons). Together these cases show that GPIs are indeed powerful, their influence is not isomorphic with material resources, and their influence operates across a wider set of actors and issues than previously demonstrated.

This symposium also adds insights into broader questions of who can exercise power and authority through GPIs. Three of our symposium contributions examine GPIs created by IGOs with varying degrees of leverage: The United Nations, while comprehensive, is often considered rather toothless; the World Bank has considerable economic leverage but claims to maintain a firewall between lending and ranking decisions; and the FATF is a standard-setting body without enforcement powers. Even governments' decisions about education spending do not appear to depend exclusively on how the MDGs affect foreign aid flows. The evidence points more towards the domestic information effects and elite bureaucratic channels of MDGs. Sometimes even poorly-resourced NGOs such as Publish What You Fund can successfully leverage GPIs. Empirical evidence across different types of actors shows that it is not necessary that the GPI creator be an already-powerful state or IGO.

Can GPIs work in the face of disagreement on what constitutes appropriate policy, or do they primarily nudge states in directions that reflect broad social consensus? Again, the articles in this symposium suggest some answers. The World Bank's focus on business deregulation has sparked massive pushback from environmental and labor groups and major powers such as

China.[59] Yet despite these dissenters, numerous governments try to climb the ranking in the belief that a poor ranking is bad for their investment climate. The MDGs' focus on education enrollment, which critics have long argued fails to capture the quality of education, is also controversial.[60] Even a topic as seemingly innocuous as aid transparency has stoked debate about whether this focus is productive. Lack of consensus often spurs GPI promulgation in the first place. Thus, consensus is not a necessary condition for GPI success, although we would expect rankings to have even more clout when they reinforce broadly held social values.

Our framework suggests some important scope conditions, which should be the subject of future research. First, domestic governing structures can be expected to produce heterogeneous outcomes. Regime type, the density of civil society actors and the information environment are all likely to temper GPI effects.[61] Second, GPI influence will also vary based on local values. Although GPIs have strong agenda-setting powers, they cannot be expected to force action to which local elites are utterly opposed or where the message does not resonate at all with local actors. Third, GPIs are limited by the capacity of the rated entities to respond. A range of conditions can undermine state capacity, from pervasive political violence to corruption to severe resource shortages. As a corollary, GPIs are more likely to be influential when they are *actionable*. It is considerably easier for all countries to change a fee, for example, than it is to quickly improve child mortality or perceptions about corruption. This likely explains the near absence of GPIs focused on military issues. The security-related indices that do exist resemble simple data without ambitious statements that they intend to influence state policies. Fourth, when the creator lacks social importance to the target, GPIs are easily ignored. Again, this

---

[59] Doshi, Kelley and Simmons this symposium.
[60] Filmer, Hasan, and Pritchett 2006.
[61] Dai 2007. See also Hall and Biersteker 2002 ; Simmons 2009. Koppell 2005.

suggests a corollary: social isolates, such as North Korea, are inherently weak candidates for GPI influence. As with many forms of external pressure, there may be a sweet spot for maximal GPI impact centering on transitioning democracies, the middle-income countries and/or emerging markets.

In addition, while it is reasonable to hypothesize that, all things equal, material leverage or predominant power increase GPI influence, the contributions show that it is not necessary for the creators to possess these traits. Our framework proposes – and the empirical research shows – that GPIs work primarily through reputational concerns sparked by comparative information. Even in the case of the FATF blacklist, which relies primarily on markets to incentivize policy change, Morse finds that officials profess to be more motivated by a desire for a clean international reputation than a specific market reaction. The symposium research thus consolidates prior claims that GPIs are strongly driven by reputational concerns while allowing that material leverage and reputation-engaging information likely interact.[62]

Finally, this symposium has called attention to the possibility of unintended effects. GPIs target specific policies and behaviors, but if governments operate with fixed resources, they may shift their priorities to meet GPI expectations, as shown by Bisbee et al. States tend to game ranking systems, something that at least some states like Georgia appear to have done with respect to the *EDB Index*.[63] This does not mean that all concerted efforts to improve are shallow, but it does highlight a possible cost of using rankings for motivation.

---

[62] Kelley and Simmons 2015 explore, but do not find, such interactive effects, while Kelley 2017 finds some evidence that material leverage enhances GPI effects.
[63] Schueth 2011.

IV.   An Agenda for Research

Globalization, accountability, and technology are changing important aspects of global governance. While coercion, enforcement and material sanctions have often taken pride of place as major movers of interstate relations, scholars and policy agents alike have come to appreciate the multifaceted nature of power exerted more subtly and gradually.[64] The proliferation of GPIs is one example of such power. They contain ideas and worldviews, and they attempt to "regulate" through non-coercive means. When promulgated by authoritative actors, GPIs may have direct, indirect and even unintended effects. Establishing causality is challenging, but the symposium contributors present corroborating evidence to establish highly plausible channels of GPI influence.

Research on GPIs raises a profound issue: how does assessment power affect worldwide authority relationships in the modern information age more generally?[65] At present it is too early to be sure. Our data show that rich states and wealthy actors overwhelmingly do the rating, while the poor are often the ones being assessed. GPIs are potentially more useful for nudging policy change than traditional coercive power resources, like economic and military sanctions. They effectively supplement other sources of leverage under the control of developed countries. Some evidence suggests their creators realize this and do not squander perceptions of GPI legitimacy with overt political or economic pressure, as the firewalling of the EDB office from the loan office at the World Bank suggests. The FATF does not sanction non-compliant jurisdictions and does not require or even urge its membership to do so. These powerful actors – who could in theory choose otherwise – prefer to stage competition and engage reputations rather than threaten

---

[64] Baldwin 2016 ; Barnett and Duvall 2005.
[65] Büthe 2012 ; Davis, Kingsbury, and Merry 2012 ; Löwenheim 2008.

sanctions or exert political pressure. Non-state actors have also used GPIs to influence discourse,

agendas and, ultimately, outcomes, and the case of *Aid Transparency Index* suggests they are

sometimes successful.[66] But to date, the authority to define the game and to keep score rather

firmly resides with existing centers of global power.

Future research should prioritize several issues. First, primary research informed by

psychology and sociology should investigate what makes ratings and rankings – and the

information on which they are based – believable. Second, more research can be done on the

information milieu in which GPIs compete. The papers in this symposium are clear about the

information ecology of their specific GPI, but it would be useful to know much more about

whether and how GPIs compete with other forms of information. Relatedly, it would be useful to

know not only why some GPIs seem to succeed, but also why others gain no traction at all. This

would require further investigation of first mover advantages, the authority of the promulgator,

and the marketing of ideas in the form of comparative data globally. Third, we envision more

intensive research on informational contestation. We have provided a flavor of that contestation,

but much more can be done to theorize the politics of information gathering and deployment

globally. Fourth, scholars should explore how non-state actors develop sophisticated information

campaigns. While shaming has taken center stage in the study of global advocacy groups, less

attention has been given to advocates' capacity to develop and deploy methodologically

transparent performance measures. Fifth, as with any policy instrument, future research should

focus not only on whether they obtain their objectives but also at what possible costs, including

possible unintended consequences. Sixth, while we have shown that an array of different types of

actors use GPIs with success, future research should test the limits of the claim that GPIs can

---

[66] Avant, Finnemore, and Sell 2010 ; Hale and Held 2011 ; Hale and Roger 2014.

disperse influence beyond traditional power centers. Seventh, our contribution has found a strong

rule of reputation, yet we also recognize the link with material leverage. Future research should

continue to investigate the *interactive* consequences of material leverage and reputation-

engaging information.

Finally, it would be useful to link research on GPIs as social knowledge and social

pressure with other new forms of global governance. GPIs are unique in their reiterated, public

and comparative nature. They are potentially useful as a contemporary 'hybrid' governance tool,

along with other methods of external influence, such as public-private partnerships,[67]

experimental governance,[68] and multi-stakeholder initiatives.[69] Now may be the time to think

creatively about how GPIs support or undermine governance in the presence of what Michael

Zürn refers to as growing global governance dilemmas.[70] As the increasing prevalence of

interstate efforts to manipulate social media and other information flows shows, strategic use of

information is rapidly becoming an ever more potent tool of influence. Addressing GPI impact

on discourse, governance and global and local power relations is an ambitious research agenda,

but the ideas in this symposium provide a strong start.

---

[67] Börzel and Risse 2005 ; Osborne 2002 ; Schäferhoff, Campe, and Kaan 2009.
[68] De Búrca, Keohane, and Sabel 2014 ; Szyszczak 2006.
[69] Bäckstrand 2006 ; O'Rourke 2006.
[70] Zürn 2018.

References:

Aaron, Joshua R., Amy McMillan, and Brandon N. Cline. 2012. Investor Reaction to Firm Environmental Management Reputation. *Corp Reputation Rev* 15 (4): 304-318.

Adair, John G. 1984. The Hawthorne effect: A reconsideration of the methodological artifact. . *Journal of Applied Psychology* 69 (2): 334-345.

Apaza, Carmen R. 2009. Measuring governance and corruption through the worldwide governance indicators: Critiques, responses, and ongoing scholarly discussion. *PS: Political Science & Politics* 42 (01): 139-143.

Arndt, Christiane. 2008. The Politics of Governance Ratings. *International Public Management Journal* 11 (3): 275-297.

Arndt, Christiane, Charles Oman, Co-operation Organisation for Economic, and Centre Development. Development. 2006. *Uses and abuses of governance indicators*. of *Development Centre studies*. Paris: Development Centre of the Organisation for Economic Co-operation and Development.

Ashmore, Richard D, Kay Deaux, and Tracy McLaughlin-Volpe. 2004. An Organizing Framework for Collective Identity: Articulation and Significance of Multidimensionalit. *Psychological Bulletin* 130 (1): 80-114.

Avant, Deborah D., Martha Finnemore, and Susan K. Sell, eds. 2010. *Who governs the globe?*, *Cambridge studies in international relations ; 114*. New York: Cambridge University Press.

Bäckstrand, Karin. 2006. Multi-stakeholder partnerships for sustainable development: rethinking legitimacy, accountability and effectiveness. *European Environment* 16 (5): 290-306.

Baldwin, David A. 2016. *Power and international relations: A conceptual approach*. Princeton University Press.

Barnett, Michael, and Raymond Duvall. 2005. Power in international politics. *International organization* 59 (1): 39-75.

Blader, Steven L., and Ya-Ru Chen. 2012. Differentiating the effects of status and power: A justice perspective. *Journal of Personality and Social Psychology* 102 (5): 994-1014.

Borgatti, Stephen P, and Rob Cross. 2003. A relational view of information seeking and learning in social networks. *Management science* 49 (4): 432-445.

Börzel, Tanja, and Thomas Risse. 2005. "Public-Private Partnerships: Effective and legitimate tools of international governance." In *Complex Sovereignty*, eds. Lou Pauly and Edgar Grande. Toronto: University of Toronto Press. 195-216.

Bourdieu, Pierre. 1989. Social space and symbolic power. *Sociological theory* 7 (1): 14-25.

Broome, Andre, and Joel Quirk. 2015. Governing the world at a distance: the practice of global benchmarking. *Review of International Studies* 41 (Special Issue 05): 819-841.

Buchanan, Allen, and Robert O Keohane. 2006. The legitimacy of global governance institutions. *Ethics & international affairs* 20 (4): 405-437.

Büthe, Tim. 2012. Beyond supply and demand: a political-economic conceptual model. *Governance by indicators: global power through classification and rankings*: 29-51.

Carpenter, R Charli. 2007. Setting the advocacy agenda: Theorizing issue emergence and nonemergence in transnational advocacy networks. *International Studies Quarterly* 51 (1): 99-120.

Carpenter, R. Charli. 2011. Vetting the Advocacy Agenda: Network Centrality and the Paradox of Weapons Norms. *International Organization* 65 (01): 69-102.

Chayes, Abram, and Antonia Handler Chayes. 1995. *The new sovereignty: compliance with international regulatory agreements*. Cambridge MA: Harvard University Press.

Checkel, Jeffrey T. 2001. Why comply? Social learning and European identity change. *International Organization* 55 (3): 553-588.

Chwieroth, Jeffrey M. 2013. "The silent revolution:" How the staff exercise informal governance over IMF lending. *The Review of international organizations* 8 (2): 265-290.

Dafoe, Allan, Jonathan Renshon, and Paul Huth. 2014. Reputation and status as motives for war. *Available at SSRN 2332048*.

Dai, Xinyuan. 2007. *International institutions and national policies*. Cambridge, UK ; New York: Cambridge University Press.

Davenport, Thomas H, Robert G Eccles, and Laurence Prusak. 1998. Information politics. *The strategic management of intellectual capital*: 101-120.

Davis, Kevin E, Benedict Kingsbury, and Sally Engle Merry. 2012. Indicators as a technology of global governance. *Law & Society Review* 46 (1): 71-104.

Davis, Kevin E., Angelina Fisher, Benedict Kingsbury, and Sally Engle Merry, eds. 2012. *Governance by indicators : global power through classification and rankings*. Oxford: Oxford University Press.

Davis, Kevin E., Sally Engle Merry, and Benedict Kingsbury. 2015. *The quiet power of indicators : measuring governance, corruption, and the rule of law*. Cambridge University Press.

De Búrca, Gráinne, Robert O. Keohane, and Charles Sabel. 2014. Global Experimentalist Governance. *British Journal of Political Science* 44 (03): 477-486.

Espeland, Wendy Nelson, and Mitchell L Stevens. 1998. Commensuration as a social process. *Annual review of sociology* 24 (1): 313-343.

Espeland, Wendy Nelson, and Michael Sauder. 2007. Rankings and Reactivity: How Public Measures Recreate Social Worlds. *American Journal of Sociology* 113 (1): 1-40.

Filmer, Deon, Amer Hasan, and Lant Pritchett. 2006. A millennium learning goal: Measuring real progress in education.

Fioramonti, Lorenzo. 2013. *Gross domestic problem: The politics behind the world's most powerful number*. Zed Books.

Ginsburg, Alan, Geneise Cooke, Steve Leinwand, Jay Noell, and Elizabeth Pollock. 2005. Reassessing US International Mathematics Performance: New Findings from the 2003 TIMSS and PISA. *American Institutes for Research*.

Hale, Thomas, and David Held. 2011. *Handbook of transnational governance*. Polity.

Hale, Thomas, and Charles Roger. 2014. Orchestration and transnational climate governance. *The Review of International Organizations* 9 (1): 59-82.

Hall, Rodney Bruce, and Thomas J Biersteker. 2002. *The emergence of private authority in global governance*. Vol. 85. Cambridge University Press.

Hammersley, Martyn. 2005. Is the evidence-based practice movement doing more good than harm? *Evidence & policy* 1 (1): 85-100.

Hansen, Hans Krause. 2012. The power of performance indices in the global politics of anti-corruption. *Journal of International Relations and Development* 15 (4): 506-531.

Harsanyi, John C. 1966. A bargaining model for social status in informal groups and formal organizations. *Behavioral Science* 11 (5): 357-369.

Heiss, Andrew, and Judith G. Kelley. 2017. From the Trenches: A Global Survey of Anti-TIP NGOs and Their Views of U.S. Efforts. *Journal of Human Trafficking* 3 (3): 231-254.

Hendrix, Cullen S, and Wendy H Wong. 2013. When is the pen truly mighty? Regime type and the efficacy of naming and shaming in curbing human rights abuses. *British Journal of Political Science* 43 (03): 651-672.

Høyland, Bjørn, Karl Moene, and Fredrik Willumsen. 2012. The tyranny of international index rankings. *Journal of Development Economics* 97 (1): 1-14.

Joachim, Jutta M., Bob Reinalda, and Bertjan Verbeek, eds. 2008. *International organizations and implementation : enforcers, managers, authorities?* London ; New York: Routledge/ECPR.

Johnson, Matthew. 2017. "Regulation by Shaming: Deterrence Effects of Publicizing Violations of Workplace Safety and Health Laws." Duke University.

Johnston, Alastair Iain. 2001. Treating International Institutions as Social Environments. *International Studies Quarterly* 45 (4): 487-516.

Karpik, Lucien. 2010. *Valuing the unique : the economics of singularities*. Princeton: Princeton University Press.

Kaufmann, Daniel, Aart Kraay, and Pablo Zoido. August 1999. Governance Matters. *World Bank Policy Research Working Paper No. 2196*: Available at SSRN: https://ssrn.com/abstract=188568.

Keck, Margaret E., and Kathryn Sikkink. 1998. *Activists beyond borders : advocacy networks in international politics*. Ithaca, N.Y.: Cornell University Press.

Kelley, Judith G. 2017. *Scorecard diplomacy: Grading states to influence their reputation and behavior*. Cambridge University Press.

Koppell, Jonathan GS. 2005. Pathologies of accountability: ICANN and the challenge of "multiple accountabilities disorder". *Public administration review* 65 (1): 94-108.

Larson, James R., and Christine Callahan. 1990. Performance monitoring: How it affects work productivity. *Journal of Applied Psychology* 75 (5): 530-538.

Latané, Bibb 1981. American Psychologist. *The psychology of social impact* 36 (4): 343-356.

Lebovic, James H, and Erik Voeten. 2006. The Politics of Shame: The Condemnation of Country Human Rights Practices in the UNCHR. *International Studies Quarterly* 50 (4): 861-888.

Levy, Marc A. 1993. "European Acid Rain: The Power of Tote-Board Diplomacy." In *Institutions for the earth: sources of effective international environmental protection*, eds. Peter M Haas, Robert Owen Keohane and Marc A Levy. Cambridge: Mit Press. 75-132.

Löwenheim, Oded. 2008. Examining the State: a Foucauldian perspective on international 'governance indicators'. *Third World Quarterly* 29 (2): 255-274.

Luca, Michael, and Jonathan Smith. 2013. Salience in Quality Disclosure: Evidence from the U.S. News College Rankings. *Journal of Economics & Management Strategy* 22 (1): 58-77.

Malito, Debora Valentina , Gaby Umbach, and Nehal Bhuta, eds. 2018. *The Palgrave Handbook of Indicators in Global Governance*. Cham: Palgrave MacMillan.

Manners, Ian. 2002. Normative power Europe: a contradiction in terms? *JCMS: Journal of common market studies* 40 (2): 235-258.

McCombs, Maxwell E, and Donald L Shaw. 1972. The agenda-setting function of mass media. *Public opinion quarterly* 36 (2): 176-187.

Meredith, Marc. 2004. Why Do Universities Compete in the Ratings Game? An Empirical Analysis of the Effects of the U.S. News and World Report College Rankings. *Research in Higher Education* 45 (5): 443-461.

Merry, Sally Engle, and John M Conley. 2011. Measuring the world: indicators, human rights, and global governance. *Current Anthropology* 52 (S3): 000-000.

Miller, Clark A. 2005. New Civic Epistemologies of Quantification: Making Sense of Indicators of Local and Global Sustainability. *Science, Technology & Human Values* 30 (3): 403-432.

Monks, James, and Ronald G. Ehrenberg. 1999. U.S. News & World Report's College Rankings: Why They Do Matter. *Change: The Magazine of Higher Learning* 31 (6): 42-51.

Murdie, Amanda M., and David R. Davis. 2012. Shaming and Blaming: Using Events Data to Assess the Impact of Human Rights INGOs1. *International Studies Quarterly* 56 (1): 1-16.

Murguia, Juan M, and Sergio H Lence. 2014. Investors' Reaction to Environmental Performance: A Global Perspective of the Newsweek 's "Green Rankings". *Environmental and Resource Economics*: 1-23.

Nelson, Stephen Craig. 2017. *The Currency of Confidence: How Economic Beliefs Shape the IMF's Relationship with Its Borrowers*. Cornell University Press.

Nutley, Sandra M, Huw TO Davies, and Peter C Smith. 2000. *What works?: Evidence-based policy and practice in public services*. MIT Press.

O'Rourke, Dara. 2006. Multi-stakeholder regulation: privatizing or socializing global labor standards? *World development* 34 (5): 899-918.

Osborne, Stephen. 2002. *Public-private partnerships: Theory and practice in international perspective*. Routledge.

Parks, Bradley, and Takaaki Masaki. 2016. When Do Performance Assessments Influence Policy Behavior? Micro-Evidence from the 2014 Reform Efforts Survey. *Presented at the APSA Annual Conference, Mini-Conference on Assessment Power in World Politics, Philadelphia PA*.

Pawson, Ray. 2006. *Evidence-based policy: A realist perspective*. Sage publications.

Pope, Devin G. 2009. Reacting to rankings: Evidence from "America's Best Hospitals". *Journal of Health Economics* 28 (6): 1154-1165.

Rege, Mari. 2008. Why do people care about social status? *Journal of Economic Behavior & Organization* 66 (2): 233-242.

Ridgeway, Cecilia L. 2013. Why Status Matters for Inequality. *American Sociological Review*.

Rieh, Soo Young. 2002. Judgment of information quality and cognitive authority in the Web. *Journal of the American Society for Information Science and Technology* 53 (2): 145-161.

Risse, Thomas. 1999. International Norms and Domestic Change: Arguing and Communicative Behavior in the Human Rights Area. *Politics and Society* 27 (4): 529-559.

Ron, James, Howard Ramos, and Kathleen Rodgers. 2005. Transnational Information Politics: NGO Human Rights Reporting, 1986–2000. *International Studies Quarterly* 49 (3): 557-587.

Rosen, Stephen Peter. 2005. *War and human nature*. Princeton, N.J.: Princeton University Press.

Sanderson, Ian. 2002. Evaluation, policy learning and evidence-based policy making. *Public administration* 80 (1): 1-22.

Schäferhoff, Marco, Sabine Campe, and Christopher Kaan. 2009. Transnational Public-Private Partnerships in International Relations. *International Studies Review* 11 (3): 451-474.

Scheufele, Dietram, and David Tewksbury. 2007. Framing, Agenda Setting, and Priming: The Evolution of Three Media Effects Models. *Journal of Communication* 57 (1): 9-20.

Schueth, Sam. 2011. Assembling international competitiveness: The Republic of Georgia, USAID, and the Doing Business project. *Economic geography* 87 (1): 51-77.

Sikkink, Kathryn. 1993. Human Rights, Principled Issue-Networks, and Sovereignty in Latin America. *International Organization* 47 (3): 411-441.

Simmons, Beth A. 2009. *Mobilizing for human rights : international law in domestic politics*. New York: Cambridge University Press.

Simonsohn, Uri. 2011. Lessons from an "Oops" at Consumer Reports: Consumers Follow Experts and Ignore Invalid Information. *Journal of Marketing Research* 48 (1): 1-12.

Sinclair, Timothy J. 2008. *The new masters of capital: American bond rating agencies and the politics of creditworthiness*. Cornell University Press.

Stanton, Elizabeth A. 2007. The human development index: A history.

Stone, Diane. 2002. Introduction: global knowledge and advocacy networks. *Global Networks* 2 (1): 1-12.

Szyszczak, Erika. 2006. Experimental governance: the open method of coordination. *European Law Journal* 12 (4): 486-502.

Thomas, Melissa A. 2010. What do the worldwide governance indicators measure? *The European Journal of Development Research* 22 (1): 31-54.

Tingley, Dustin H., and Barbara F. Walter. 2011. The Effect of Repeated Play on Reputation Building: An Experimental Approach. *International Organization* 65 (2): 343-365.

van de Walle, Steven. 2006. The state of the world's bureaucracies. *Journal of Comparative Policy Analysis: Research and Practice* 8 (4): 437-448.

van Ham, Peter. 2001. The Rise of the Brand State: The Postmodern Politics of Image and Reputation. *Foreign Affairs* 80 (5): 2-6.

Weisband, Edward. 2000. Discursive Multilateralism: Global benchmarks, Shame, and Learning in the ILO Labor Standards Monitoring Regime. *International Studies Quarterly* 44 (4): 643-666.

Wilson, Patrick. 1983. *Second-hand knowledge : an inquiry into cognitive authority*. Westport, Conn.: Greenwood Press.

Wohlforth, William, C. 2009. Unipolarity, Status Competition, and Great Power War. *World Politics* 61 (1): 28-57.

Zürn, Michael. 2004. Global governance and legitimacy problems. *Government and Opposition* 39 (2): 260-287.

———. 2018. *A Theory of Global Governance: Authority, Legitimacy, and Contestation*. Oxford University Press.