Knowledge Packaging Technologies: Organizing Status Competition in Global Governance

By Paul Beaumont
Norwegian University of Life Sciences
Email: paul.beaumont@nmbu.no
Telephone: 47+ 46777644
Abstract

Since the 1990s the number of organizations that rank countries according to performance in social indicators has burgeoned. Yet, the political effects of rankings are poorly understood. To date, most research tends to investigate their methodological veracity rather than how to understand their influence on the social world. Addressing this lacuna, this paper foregrounds the political consequences of these rankings by theorizing how presenting knowledge in rankings can engender a peculiar type of “status shock” that provides the discursive resources to legitimize policy reform. It is argued that rankings may operate as a status recognition conduit, whereby states grant it authority to allocate status in given policy fields. Problematizing status among peers in a given social hierarchy, rankings can, under some circumstances, set in motion a continuous process of status competition among the rankees. By way of illustration, the paper uses Norway’s response to the PISA education ranking to explore the mechanism through which these technologies affect their target population. The case reveals that the publication of PISA rankings in Norway induced a “status shock”, necessary to legitimate reforms of education.

Key Words: Status seeking, hierarchy, global governance, norms, institutions, international organizations
Since the 1990s the number of global ranking organizations that rank countries according to performance in social indicators has exploded since the 1990s. From gender equality, to corruption, from ease of doing business to trafficking, these rankings claim to make visible states' relative performance in all manner of social fields. (Cooley and Snyder, 2015). Yet, the political effects of rankings are poorly understood. To date, most research has tended to focus on their methodological veracity rather than their influence on the very subjects they rank (Cooley & Snyder, 2015). Recently this has begun to be addressed, though theorizing about the political, social and ethical consequences of this new “technology” of governance is in a preliminary stage (Jerven, 2016). While liberal institutionalism illuminates the role of organizations providing information and reducing transactions costs to facilitate co-operation, this only loosely captures rankings, which provide information and reduce transaction costs seemingly to facilitate competition. Indeed, many of the rankings measure excludable, non-rival goods – one can pursue gender equality, education, etc. without affecting others’ ability to do the same — and therefore the social qualities ranked rarely resemble the “games” found in global public goods theorizing (e.g. Betts, 2009). Further, liberal institutionalism offers only limited purchase on why actors would compete for position when leading or lagging offers little direct material reward or punishment. Indeed, unlike EU law, which operates under legal-rational authority, or the World Bank, which can offer economic sticks and carrots tied to its structural adjustment packages, many global rankings have no such means to influence policy. Constructivist norm scholars should be well-placed to fill this gap, however as yet they have given little attention to the new governmental “technology” of rankings (Miller and Rose, 2008), specifically, to the mechanism through which they may induce policy change (Jerven, 2016). This then begs the question, lacking legal authority and economic incentives, whether, how, and when, international rankings influence government policies?

Exploring this question, this paper foregrounds the political consequences of these rankings by bringing together the burgeoning status-seeking research agenda,¹ and the nascent

research into global rankings. These literatures should make natural bedfellows: rankings by definition put competitors relative position in social hierarchy into public domain, meanwhile status research investigates how an actor’s concern for position in a social hierarchy informs their policies. Yet curiously, there is negligible cross-fertilization to be found in the respective literatures. For instance, while Merry et al (2015) mention in their introductory chapter that indicators “confer status on countries ranked highly and stigmatizes those ranked lower” none of the 17 empirical chapters take up this line of enquiry. Similarly, Cooley and Snyder (2015) introductory chapter in the edited volume theorize that rankings “might exert influence by impacting the recipient’s social status” but do not cite status research in IR, nor do the empirical chapters follow up this line of inquiry. Meanwhile, the seminal edited volume in contemporary status seeking research (Paul, et al, 2014), makes just one passing reference to global rankings.

Thus, building a much-needed bridge between these camps, this paper theorizes how and when rankings may induce change inspired by status concerns. Following Miller and Rose’s, governmentality frame, this article conceptualizes rankings a new type of knowledge packaging technology, which absent material incentives or legal authority, induce change by illuminating and thus problematizing states’ performance among peers in a given policy field. I argue that under certain conditions rankings may induce status shocks: public opprobrium among the public and elites that their country compares poorly to their rivals and worse than the public had expected. This may then be used as a reason to reform the policy in question with the aim of rectifying their poor ranking position. Over a longer period, rankings create a


3 Irving Goffman-inspired stigma research in their introduction but overlook all major status seeking research of the past decade
“playing field” for a status competition. Because rankings are relative rather than based on absolute standards, provided the government values the game, rankings can offer continuous potential reason to reform (for all participants not placing first). As this suggests, the process is not automatic; it is necessary to theorize how rankings generate responses by multiple domestic policy actors (Cooley 2015). This article suggests that rankings provide discursive resources to domestic actors that can enable them to legitimize policy reforms that would otherwise face more domestic opposition, or vice versa: provide ammunition for opponents to criticize the government and engender reform. Such a theorization makes a further novel contribution to the status literature, which almost universally assumes states generate status by direct practices of recognition from other states. In contrast, this article theorizes global rankings – like the OECD’s Programme for International Student Assessment (PISA)— may operate as a status recognition conduits: taking on the role of a governing body that assesses and allocates status recognition for good and bad performance.

By way of probing the plausibility (Eckstein, 2000) of whether global rankings could affect change in this manner, this article uses Norway’s responses to the PISA education ranking. PISA was selected as an easy case for exploring whether rankings can induce status shocks and/or status competition. As I will outline below, the OECD’s authority in education, PISA’s global dissemination, and the fact that its referent—high quality education—is a valued attribute in all modern societies, implies that if PISA rankings did not inspire status competition here then rankings would be unlikely to anywhere. The case vignette does indeed strongly indicate that the publication of PISA rankings in Norway induced a “status shock”, that helped legitimate significant reforms of education. Furthermore, a competitive dynamic also seems to have emerged in a manner congruent with status competition. Ultimately, the paper suggests this tentative exploration provides more than enough evidence to further investigate when, how, why and packaging knowledge in rankings induce status shocks and status competition.

This paper proceeds in four moves. The first situates rankings within the extant literature on global governance and argues that the liberal institutionalist and weak constructivist literatures lacks the analytical tools to make sense of rankings social power. Instead, I suggest framing rankings as governmental technology, that without sticks or carrots renders technical

---

4 The three major edited volumes on status recognition
and problematizes state’s status in policy fields in order to exert social pressure upon the rankees and encourage policy reform. The next section draws upon status theory from social psychology explains how status concerns can influence behavior and why theoretically rankings could induce status concerns. The third section outlines the specific mechanism through which I argue rankings can influence states policy. The final section uses PISA and Norway’s response to PISA to illustrate this mechanism.

**Governing Anarchy by Rankings**

To begin, global rankings must be situated within the literature on how global governance is said to work (or not). Without belaboring the obvious, whether it is possible, and if so how, governing sovereign states in “anarchy” remains a central problematic in international relations and International Relations respectively. Maybe the promise of institutions is false, but whatever way one cuts it, international organizations and the global governance initiatives they manifest, usually feature in the problem, answer and/or the solution. The extant literature on how institutions foster change tends to coagulate along the poles in the material/ideational divide (Bauhr and Nasiritousi, 2012: 543). Indeed, while the prospects for institutions to independently alter the material incentives facing states and facilitate cooperation constituted the mainstay of the 1990s neo-neo debate (e.g.Keohane and Martin, 1995), constructivist scholarship focused on how institutions could facilitate the spread of new norms via learning and persuasion, which if successful would ultimately lead to internalization and semi-automatic norms adherence (e.g.Finnemore and Sikkink, 1998; Wendt, 1999). It is not necessary for our purposes to contest the relative merits of these theories of how institutions work other than to explain how and why they struggle to adequately capture how rankings may influence the ranked.

Neither changing the payoffs nor persuading an actor of their normative failings, exhaust the reasons why states may respond to institutions and especially rankings (Johnstone, 2001; Axelrod, 1997). While it is fairly self-evident why rankings divorced from sticks or carrots sits ill with liberal institutionalists, so we will focus here on the other of Bauhr and Nasiritousi’s poles: ideational approaches that emphasize the importance of dialogue and persuasion in “socializing” a target population to “internalize” a given norm. First, states
may undertake “pro social behavior” like seeking to improve their position in a ranking, without necessarily internalizing the norm the ranking embodies (Johnstone, 2001; Zarakol, 2014). Given its prevalence in inter-state discourse, Johnstone suggests the “most important microprocess of social influence (...) is the desire to maximize status, honor, prestige-diffuse reputation or image-and the desire to avoid a loss of status, shaming, or humiliation.”

Critically, none of these desires requires the internalization of a norm, but only that an actor values their social status and realizes that adherence to the norm depends upon following it. Second, rankings do does not fit very easily with the notion of persuasion as it is usually understood by weak constructivists. The process of determining the methodologies of rankings is usually opaque, technocratic, and once created, rankings require high degree of expertise and the social power to contest them (Merry, 2011; Cooley & Snyder, 2015). Failing these unlikely conditions, governments are left with choice of rejecting, ignoring the rankings, or accepting and perhaps seeking to improve performance in. None of these options remotely resemble the ideal of a Habermasian dialogue, constructivist notions of persuasion are based upon. Finally, rankings induce a different kind of relationship between the actors: instead of trying of trying to become “like” the others as weak constructivists envisage, the participants seek to best one another. Thus the dynamic changes from one of homogenization and learning to one of stratification and social competition (Towns and Rumelili, 2017).

Having established how mainstream work on institutions offers little purchase on rankings potential to exert influence, we need a means of conceptualizing how organizations lacking sticks, carrots, or persuasive argument can influence sovereign states. Michel Foucault’s conceptualization of government—“the conduct of conduct”—offers a productive (not to mention popular) way to illuminate the effects global governance, even when the governments themselves remain sovereign. A governmentality lens allows analysis of governmental practices without privileging the sovereign state subject as locus of political power (Miller and Rose, 2008). Therefore, a governmentality lens enables an analysis of processes and techniques of government rather than monadic entities (like the state) undertaking discrete actions upon their subjects (see Jackson and Nexon, 1999). For Foucault, the unitary instrumental state is a “grand myth” (Foucault, 1991) instead, he directs us to the practices of “disposing of things”: forms “of activity to shape, guide or affect the conduct of person or persons” (Gordon, 1991: 3). The upshot of Foucault’s definition of government vis a vis the sovereign state, is that it permits analyzing the emergence global governance without the pre-requisite of a world government wielding legitimate authority to
govern first (in the Weberian sense). Instead, it illuminates the governance technologies— for example, the nascent practice of ranking—which while lacking support from a monopoly on violence, nonetheless undertake forms of activity aimed to *shape, guide, or affect* the conduct relations between persons and things, and/or collectivities. In this optic, the burgeoning number of international ranking organizations become visible, and can thus be conceived of as a new *governmental technology*. This packaging technology seeks to manage populations by strategically making visible relative *ranking of groups* potentially fermenting competitive dynamics between those groups.

With this idea in hand, we can turn to the technologies of government that are at the disposal of transnational actors. In her analysis of NGO’s shortcomings in Indonesia, Tanya Li’s discussion of governmental technologies provides a useful framework for understanding two inter-related procedures for governing populations. First, *problematization*: rendering visible problems to be rectified. The second process Li conceptualizes as *rendering technical*: which involves defining boundaries, acquiring relevant information and techniques for correction. (Li, 2007: 6). These two processes seek to “foster beneficial processes and mitigate destructive ones. They may operate on population in the aggregate, or on subgroups divided by gender, location, age, income, or race, each with characteristic deficiencies that serve as points of entry for corrective interventions.” (Reference) Ranking organizations fit Li’s conceptualization of governmental technologies well: they seek to problematize behavior by rendering technical and publicizing a deficiency of a population (low status); and imply a solution within specific boundaries defined by their indicators methodology (the values embedded within them). But what is the effect that these rankings seek to induce on their target population?

**Global Rankings and Status Research**

---

This paper is interested in the novel way in which these ranking seek to exert social influence: by inciting status concern and competition among states. To be clear, rankings can exert pressure in other ways, most obviously when they are linked directly to a material carrot or stick. It is this, non-materilly incentivized pressure that this paper seeks to investigate. In order to theorize how global rankings could inspire status shocks and potentially status competition, it is necessary to specify our definition. Davis et al. (2012: 5) provide a suitably conventional definition of global indicators:

An indicator is a named collection of rank-ordered data that purports to represent the past or projected performance of different units. The data are generated through a process that simplifies raw data about a complex social phenomenon. The data, in this simplified and processed form, are capable of being used to compare particular units of analysis (such as countries or institutions or corporations), synchronically or over time, and to evaluate their performance by reference to one or more standards. (my emphasis)

I have highlighted the key components of this definition of indicators that are theoretically relevant to status competition. First, for something to have high status it requires the existence of low status, thus, it requires the existence of a hierarchy (De Carvalho and Neumann, 2015a). When indicators are formalized into transitive rankings they represent almost an idealized context for a status competition: rankings provide the structure— or playing field— within which status competition between actors can unfold. Second, for status competition to occur the status attributes must be changeable: if they cannot change (e.g. like a biologically-defined racial hierarchy) actors cannot compete, or indeed have a “performance” (Beaumont, 2016). Social indicators all fit with this; they measure social qualities that to varying degrees can change. Third, a status is always relational, one cannot know one’s status without reference to the status of others (Renshon, 2016; De Carvalho and Neumann, 2015a). Thus status performance is always based upon “global comparison” between competitors rather than internal comparison based upon an actor’s individual performance (Onuf, 1989). Indeed, with rankings one cannot know one’s position without reference to others’: One can improve performance in absolute terms, but still fall down the rankings if the others perform better. Finally, for actors to suffer status shock, or engage in status competition, the hierarchy in question must be publicly known, the quality measured must be considered socially valuable, and the ranker must be considered legitimate and credible (Kelley & Simmons 2015). These conditions are the hardest for rankings to fulfil, but they are almost always the explicit goal of ranking organizations. For instance, the
Transparency International go to great lengths to publicize its corruption rankings, present its methodology as credible, and asserts that its goal is to raise the salience of corruption globally. All the global rankings (see appendix) by definition share the goal of publicizing both the ranking and the phenomenon it ranks. In sum, rankings mechanisms seem well-designed to ferment status competition.

Given the excellent fit between what global rankings seek to do, and what conditions are required for status competition to emerge, it is odd that none of the major works on global rankings (see edited volumes by Cooley and Snyder, 2015; Davis et al., 2012) develop the link with status research and vice versa. While status research briefly emerged in the 1960s (e.g. Galtung, 1964) it lurked in the shadows of IR until Wohlforth (2009) kick-started it in the 2000s. The vast majority of this contemporary status research takes its cue from social psychology (Pouliot, 2014) and it can be roughly categorized grouped into large N research and case study research. Simplifying greatly, the large N research relies on hierarchies of material power and hierarchies of diplomatic recognition to investigate whether “discrepancies” or “inconsistencies” correlate with particular discrete foreign policy actions, usually some function of war (Renshon, 2016; Renshon, 2017; Volgy et al., 2014). The qualitative research uses case studies to suggest how seemingly unusual foreign policy behaviour (e.g. Glasnost, Brazil’s African foreign policy.) can be explained by status seeking rather than conventional security or economic explanations (e.g. Larson and Shevchenko, 2003: respectively; Stolte, 2015).

These approaches have much going for them, but both are methodologically individualist, and tend to miss much of the inter-subjective relational aspects that status competition involves. Indeed, these approaches conceive of status seeking as an individual response to some structural stimulus. While the structure in both cases is theorized as a hierarchy, the relationship to specific others within the hierarchy are bracketed. For instance, it is the general amount of recognition that a state receives that functions as the independent variable, rather than their relationship to its rivals. In contrast global rankings produce public, transitive ordering of actors, explicitly putting the relationship between competitors in the foreground. Second, the hierarchies used in Large N studies are constructed by the analyst rather than being inter-subjectively present in the “real world”. For instance, in their pioneering research into how status concerns may prompt war, Renshon (2015; 2017) and Volgy et al (2011; 2014) construct hierarchies based upon their assumptions about what
constitutes status attributes and what constitutes status recognition. In both cases they assume military power to be a key status attribute and that diplomatic recognition makes a good proxy for status recognition. While justifiable and useful for their purposes, neither exhausts the range of attributes or qualities states may rest their self-esteem upon (as Renshon, readily admits, 2017; p 52-53), while diplomatic recognition in particular seems quite far removed from the everyday practices of status signaling that governments and their populations pay attention to. In contrast, rather than being a theoretical construct used by the analyst, global rankings are inter-subjectively known and potentially *used* by political agents themselves. This has an important methodological and theoretical implication for how to investigate rankings. Because rankings are ontologically meaningful to states, rather than just (potentially) epistemologically useful to researchers, we can theorize how the rankings enter the public sphere and how they may be used once there: we can theorize them as consciously utilized, governmental technologies. We will come to this later.

Conversely, global governance research dealing with rankings is only just emerging, yet already covers a lot of ground (Cooley and Snyder, 2015; Merry et al., 2015). Painting in broad strokes, the research so far has emphasized investigating the methodological veracity of the indicators, questioned the ethical desirability of rankings, and investigating whether and how rankings affect policy. It is in this latter agenda that a gap exists to be filled by theorizing status competition as one mechanism through which rankings can foster change. As noted above the research on social effects of indicators tends to mention status without foregrounding it in their analysis. The closest of the rankings literature to do so is Kelley and Simmons (2015), which investigate the social pressure of indicators via the U.S. Governments Trafficking Annual Report (TAR). Here, the authors draw on status research to explain why some rankers will have more credibility and legitimacy than others and how this will influence their potential to influence policy. This is reasonable, however oddly they overlook how status theory can also illuminate the effect upon the rankees too. Further, while the authors mention status in their conclusion, they do not theorize it, nor is their choice of case a good fit for investigating status competition. The problem is that TAR is packaged according to absolute tiers, rather than transitive ranking. This means theoretically it would foster a playing field for a different dynamic: tier systems operate as a club good that lead to homogenization and “one off” policy change rather than stratification, and continuous competition over status (Towns & Rumelili, 2017). This is borne out by Kelly & Simmons’ methodology, which seeks to count policy change, as it progresses towards an absolute
standard, which once reached cannot be surpassed. In short, there is a limit to the social influence envisioned by both TAR and the authors, and all rankees can end on the same tier. A transitive ranking has no such limit and as such has the theoretical possibility for endless status competition. Finally, their choice of case, TAR, is also problematic for assessing the “social pressure” of indicators because the producer of the ranking, the U.S., possesses latent sticks and carrots. Thus, the authors cannot rule out these lurking powers as the source of the policy change they seek to explain. The OECD’s PISA ranking has no such lurking sources of influence.

How Rankings Work: Status Shocks, Competitions and Conduits

Social psychology literature since the late 1970s has demonstrated that individuals generate self-esteem by seeking to achieve favorable comparisons between their respective in-group and other groups (Hymans, 2002; Larson and Shevchenko, 2003). Under laboratory conditions social-psychologists have shown how individuals undertake otherwise odd behavior—e.g. going against one’s own self-interest—in order to improve the in-group’s position vis-à-vis another group (Tajfel, 1978). This phenomenon—termed Social Identity Theory (SIT)—has been translated into IR in the last decade. A recent review of this literature suggested that if nothing else, there now is convincing evidence and consensus among the literature that statesmen and national publics value (and even “obsess” over) their social group’s status (Dafoe et al., 2014). This article suggests rankings should therefore be understood as an institutionalized method of establishing a group’s status in a given field. Moreover, given that rankings are pretty much defined as a tool for allowing inter-group comparison, I suggest it is this desire felt by individuals for favorable inter-group comparisons that SIT research isolates and global rankings organizations seek to tap into.

But how could rankings affect those they rank in practice? The following section draws on the extant literature to introduce some conceptual apparatus for understanding rankings in terms of status theory, the mechanisms through which they may work, and some methodological suggestions for how we can know when they are working in this manner.

---

Two reasons are usually given for why: either intrinsic or instrumental reasons: Intrinsic implies states value status for its own sake; staking their collective self-esteem upon their status in a given hierarchy (e.g. football fans care very much about coming first, despite the lack of material reward). Instrumental status seeking implies that status in a field is useful for some other end (e.g. status as a "superpower" may reduce costs by inducing deference).
Status Shocks

Typically, in the status literature, status concerns develop when states snubbing states that believed they warranted deference (e.g. Wohlfarth, 2009; Ringmar, 2002), or more gradually, when states rise in material capability, yet do not see their privileges in international society increase accordingly (e.g. Paul, 2014, Deng, 2008). Building on this research, I suggest that rankings can construct conditions to induce status concerns in the form of a “status shock”.

As the name implies, status shocks involve the collective opprobrium that occurs when a collective learns abruptly that their status – their position in a given social hierarchy – is lower than they previously thought it to be. Thus, there needs to be credible new information about the status of the collective in a given hierarchy and this new information needs to diverge to some significant degree from what the collective believed its status was. They key here is the gap between expected recognition and received recognition. It is thus not finishing low itself that would cause a shock, but lower than expected or finishing lower than rivals.

New rankings are well suited to induce this type of status shock because they often measure and render equivalent previously opaque social qualities that would not otherwise be easily comparable e.g. press freedom, gender equality and education. Before the rankings emerge it is quite possible for several countries to share the belief that they are the leader in a given field and stake their self-esteem upon this belief. In education for example, it seems as though Switzerland, Germany, Norway and Denmark all believed their education to be superior to their peers (see below). Conversely, lots of countries finish low in lots of indicators and it may not induce status concerns. For example, if Saudi Arabia discover they rank low on gender equality it would be unlikely to cause a status shock because they would expect, and indeed perhaps want to place low. Critically, I suggest this new information need not come directly from other states signals of recognition or rejection, as the extant status research presumes, rather rankings could provide new credible information regarding a collectives’ positions in a ranking: acting as a status recognition conduit.

Status shocks do not automatically induce policy change. Rather I suggest there is a degree of contingency: states may seek to make policies to redress their status deficit, the collective may realign their expectations to mediocrity, or they may reject the ranking. When they choose the first, the rankee may find themselves engaged in a status competition.

Status competition
While rankings illuminate status positions, their effects may be analytically separated between inducing a status shock and a longer-term process of status competition. Towns (2012; 2010) work on normative change offers avenue for theorizing how rankings may induce status competition. According to Towns, norm diffusion always stratifies (as well as homogenizes) and it is this stratification that (re)produces hierarchies and stimulates change. Towns theorizes and convincingly illustrates, via historical cases of changes in gender norms, that states adopt policies in order to move up the social hierarchy, not necessarily because of material incentives or moral persuasion. In such a model, normative change creates new rules for the hierarchy that recursively then ferments further change as actors face pressure to meet the new norms. However, Towns work conceives of looser informal hierarchies: modern vs traditional, or “club goods”, rather than the transitive rankings scrutinized here. Indeed, competition in Towns’ hierarchy is not a zero-sum gain. For instance, a state can become recognized as modern without another necessarily moving back to traditional. Building on Towns, this paper ventures that rankings organizations institutionalize a hyper competitive version of this hierarchy inspired social pressure. Indeed, transitive rankings (rather than club-good hierarchies) imply that when one moves up another must move down. Moreover, it is possible to improve absolute level of performance in a given quality and still move down a transitive ranking if others improve more. Thus, when performance is captured using such a ranking, if those being ranked value their position and seek to improve it, it implies mutual competition that changes the standards for “normal”, and makes just maintaining one’s position in the rankings a competitive process. Such a process in the extreme may end up resembling Lewis Carrol’s Queen’s Race, in which “it takes all the running you can do, to keep in the same place”. If players get “taken by the game”, then it implies that rankings will provide a constant supply of discursive resources to legitimate policy reform (see below).

Institutions as Status Recognition Conduits

Most of the status literature tacitly assumes that states generate status by direct practices of recognition from individual other states. I would suggest rankings organizations can provide what I term a “status recognition conduit”. Instead of having all countries try to assess one another’s social qualities and affording and holding back status accordingly, they may recognize an organization as – such as the OECD’s PISA education rankings our case - as

7 The three major edited volumes on status recognition
legitimate arbiter of status in a given field. In short, they provide information about education that would be much costlier for states to acquire themselves. For example, pre-PISA tests, if comparisons between countries education existed it was speculative and/or anecdotal. By contrast, rankings attempt to standardize global discourse on whatever social qualities they rank. However, the success of a given ranking to function as a recognition conduit depends entirely upon its legitimacy (Kelley & Simmons 2015). Put simply, merely ranking a group of countries will not by itself induce status concerns. Rankings Organizations that are perceived as biased or lacking methodological rigour will lack legitimacy and thus struggle to have a status effect upon its rankees. For instance, few would pay attention to North Korean-produced Human Rights ranking. Conversely, countries actually pay to play in PISA. Thus it is very difficult for states to then reject their score after the fact. Finally, a private ranking may be methodologically sound but cannot induce status concerns because they need to be mobilized within national discourses in order to have a status effect.

**How Rankings Inform Policy: Reactivity and Discursive resources**

Since Hawthorn scientists have known that human subjects react to being observed and evaluated. In mainstream social science this reflexive “reactivity” is often understood as a problem to be minimized (Espeland and Sauder, 2007), but it is exactly this social effect rankings organizations seek to harness and this paper investigates. It is worth recalling Li’s formulation that governmental technologies seek to problematize a practice and render it technical: Here, previously private practices are rendered technical by measurement practices that generate equivalence, and rendered problematic via their packaging in rankings in which their comparative performance is illuminated. Knowledge of others performance relative to one’s own becomes a problem, not the absolute performance itself. It is this emphasis relative performance, which offers the novelty in this governance technology. However, even then the ranking must be filtered through domestic politics to influence policy. Rather than directly causing an effect, rankings can be understood as providing discursive resources: potentially salient information that may be mobilized in the domestic sphere to legitimate policy change or attack political opponents. This then begs the question; how can we know

---

8 It is worth noting that Miller & Rose (2008) list several related “technologies” related to auditing and indicators but although they explicitly leave the door open for rankings, they do not mention them.
when rankings have been used in this way? How can we know they were not peripheral to a policy change? In short, whether the discursive resource provided by the ranking induced action undertaken to improve position in the rankings that would not have otherwise occurred.

Drawing on the discussion above, evidence that would indicate status-driven policy would be the following two sorts of evidence in combination. First, policies that included statements that stress poor performance compared to rivals as a legitimating reason to take action. The goal is not to read minds, but rather to interpret whether a ranking made possible, or provided the discursive resources that “enabled” policy change (Neumann, 2008). Ultimately the goal is to interpret whether ranking scores were necessary for legitimizing policy change and delegitimizing inaction. The second type of evidence required, is that the policies change in accordance with the values embedded within the rankings. All rankings embed value judgements about what matters. For instance, the World Bank’s Ease of Doing Business index (EDB) values different things to the Global Competitive Index (GCI), despite purporting to measure similar things. Thus, when Georgia aimed to move up the EDB it targeted specific policies areas valued by EDB, it could end up setting a record for the fastest rise in EDB while staying more or less stationary in the GCI (Schueth, 2015). This offers an extreme illustration of how the type of policy enacted can tell us whether the policy reflected the values of the rankings. A final methodological point is that the goal here is to illuminate the mechanism by which PISA rankings can influence policy making, rather than how they must: we would not expect all countries to respond in the same way to rankings (Wight et al., 2013)

**PISA: Institutionalizing Status Competition in Education**

---

9 If the policies changed in accordance with the rankings without reference in domestic discourse then it may just be that the rankings reflect a wider discourse around that policy area. Meanwhile, if the rankings are cited but the policies do reflect the values then they may have been necessary to legitimate change, but it is probably not the status concerns that drive the change but another priority.


11 This is not to say that the burden of proof is the success of the policy, rather that one can interpret from the policy whether it can plausibly be suggested to be aimed at the ranking.
Having established how rankings provide a theoretically strong basis for status competition to emerge, and theorized how it might occur, the next section uses OECD’s PISA as an exploratory case study to investigate its plausibility. PISA seeks to measure the educational performance of students (aged 15) across the OECD and other participating countries. Since 2000, PISA has attempted to assess scientific, reading, and mathematical literacy, and recently, problem solving abilities (www.pisa.oecd.org). Yet PISA rankings do not directly compare the actual education practices. In order to simplify these complex phenomena and render “equivalence” (Li, 2015), PISA tests 500,000 students from the participating countries and uses the numerical scores it generates as a proxy for the quality of education. A great deal of research has been dedicated to outlining methodological and ethical problems associated with this process— but here, we are interested into the response packaging knowledge in a relational ranking produces on participants of PISA. Critically, PISA packages its tri-annual studies in relational rankings form. While the absolute performance and individual case studies are also presented, the relational performance in the rankings between countries is what national press emphasizes (see below). Indeed, the PISA website encourages this relational competition. When one visits Pisa.org, the most readily available means of accessing the data provides the relative score of countries performance in the rankings and compared to the average (see fig. 1.1 & 1.2), it does not provide such easy access to the absolute levels achieved in the tests. It is perhaps not surprising then, that even reputable news organizations such as the BBC present performance in rankings form, similar to how they might present performance in a sports event (See fig. 3).

Fig. 1.1 PISA education “Data” on country “performance” -- packaged by
/www.oecd.org/pisa/data/
Fig. 1.2 Country Education “Performance” “Data” packaged by /www.oecd.org/pisa/data/

Average performance

The headline indicator for the three subject areas: science, mathematics and reading. Average performance refers to all 15-year-old students in a country/economy regardless of the school type and grade attended. Small differences between countries and over time may be statistically insignificant.
This emphasis on relative performance is one reason why PISA makes an excellent case to explore status competition, but there are several others. First, PISA rankings are not linked to any material incentives: states that perform badly suffer no punishment nor receive any aid. This means rationalist materialist explanation of how rankings may induce change lack purchase, making it easier to ascertain whether status competition is at play. Second, PISA claims to measure education quality, which it is safe to assume is valued by all political communities. This is a key condition for status competition: one cannot have mutual competition if players do not value the game. This also means that PISA rankings do not persuade in the normal sense of changing the minds about the legitimacy of a norm. The norm in question, the high value attributed to good education, is firmly established in all the countries. Second, because states pay for the privilege of taking part of PISA, they tacitly grant it legitimate authority in education. Indeed, PISA itself is widely recognized as a credible indicator of education (Carvalho, 2012: 183). Fourth, PISA is widely circulated among academics, policy makers and the national press of most participating countries. Thus, PISA satisfies the requirement of status competition for being public, inter-subjectively
recognized knowledge among participants and not just an analytical construct used by scholars (e.g. rankings in number of embassies received). Ultimately, if we do not see sign of status competition here, it seems unlikely IOs rankings would prompt status competition anywhere.

“PISA Shocks”: stimulating status competition

Although much of the research on rankings alludes to the normative pressure exerted by rankings (Merry et al., 2015) few have theorized or investigated the mechanism underpinning this normative pressure. This is also true for PISA. Although PISA has spawned a burgeoning research agenda seeking to explain variation in education across countries, an OECD working paper from 2012 laments that

[…] research on its (PISA’s) effects and use within national education reform and policy-making is scarce… the causal mechanisms whereby OECD can influence national reform initiatives with its PISA study and which national factors may have a hindering or furthering effect have not been analyzed in sufficient depth. (Breakspear, 2012: my emphasis)

Nonetheless, the following seek to show that sufficient evidence exists to suggest status shocks and competition played a role in stimulating and facilitating policy changes among some participating countries.

Norway’s “PISA Shock”

It is well established that Norway suffered what has become known as a “PISA Shock” (Breakspear, 2012; Østerud, 2016). In short, PISA shocks refer to the public outcry that accompanies the results of a PISA results. In particular, Germany, Switzerland, and Norway are usually cited as examples (Breakspear, 2012). But what exactly constitutes the shock in the PISA shock? Also, to what extent and how does it influence policy? This paper contends that the “PISA shock” constitutes a status shock: the discovery that a valued attribute or quality that a social collective considered important for their collective self-esteem and a symbol of superiority over outgroups, gets undermined or challenged. The effects a status shock would imply a sense of anguish, anxiety, and calls to rectify or take urgent action to remedy regain the lost status. Norway’s response to PISA displays these
characteristics. Again to emphasize, for it to qualify as a status shock the *problem* that is illuminated is primarily articulated in reference to being worse, or falling behind other peers, *not* concern for the absolute performance in the attribute itself.

Indeed, both the press reports and the politicians’ response to PISA indicate that what most shocked Norway was their poor position in the rankings. The first ever PISA results were published in December 2001 and Norway ranked around the OECD average in maths, science and literacy (www.pisa.oecd.org). Immediately, the results prompted consternation by both press and politicians (Østerud, 2016). In a quote that provides the ideal example of how rankings ferment status competition, the minister of education, Kristin Clement, likened the discovery of Norway’s average performance to the archetypal status competition where victory gains glory rather than greed: “Norway is a school loser, now it is well documented. It is like coming home from The Winter Olympics without a gold medal (Ramnefjell, 2001, December 5). It is worth noting here that she uses the Winter Olympics for her analogy: a competition that Norway usually excels and expects to excel in. One of Norway’s leading tabloids, reported the news similarly with the headline, “Norway is a school loser” (my translation), The article goes on to complain “Norwegian 15 year old students are just average compared with their peers in the other 31 OECD countries.” (translation). Nowhere in the article does Dagbladet make reference to the absolute performance in terms of the standards that Norwegian students reached. These examples were reflective of broader public response, for instance, Østerud (2016: 15) reflects that: “The shock of that Norwegian schools were not among the best in the world, created a situation of fear and perplexity and it was expected that someone intervene quickly and put things right” (my translation). What might appear strange though, is that Norway did not perform so badly: its scores were average among the OECD and in absolute terms, not so far away from the leader, Finland.

What seemed to cause the shock was that Norway *expected* to finish top. Indeed, before the first PISA in 2001, Norwegians tended to believe their education system to be the best in the world (Baird et al., 2011). Looking at the distribution of PISA shocks as a whole, Dobbins and Martens (2010) suggest that countries who held their education system in high esteem found even average results shocking (e.g. Germany and Switzerland), meanwhile states with

---

12 Indeed, a multi-event athletic competition is Nicholas Onuf’s paradigmatic exemplar for explaining how interest in standing in formulated (Onuf, 1989, 267)
widely acknowledged problems in education did not suffer from a “PISA shock” (e.g. the U.S). Yet these theorists do not make the link to status: failing to meet expectations can be based upon a failure to meet an internally generated standard rather than relative standards. However, status research foregrounds the relational component of self-esteem that seems to underpin these states responses to their PISA performance. Nonetheless, Dobbins and Martins argument fits neatly in with status theorizing in International Relations. For instance, Renshon (2017: 21) suggests that “status concerns” are triggered when states receive less recognition than they expected, or, when they compare poorly to “significant others” (Paul et al., 2014, Hymans, 2003). Indeed, The U.S. only began to pay attention to PISA once China opted in, and was revealed to outperform the U.S (Cite). While Renshon suggests status concerns may trigger war as a means to rectify their low position, in the context of PISA, this would appear poorly suited to the task. However, PISA does enable action: educational reform.

Indeed, there is little doubt that PISA has contributed to both the timing and shape of educational reforms in Norway. The PISA Norwegian representative estimated PISA to have been “highly influential” in on education in Norway (Breakspear, 2012). Indeed, the results of the first PISA gave a push to the creation of a new national quality assessment system (NKVS) in 2004, which included national standardized tests. The specifics of the reform reflect closely PISA’s recommendations for countries seeking to improve their education: introducing more country-wide standardized testing (OECD, 2009, p. 14). Meanwhile, recalling the PISA ranking of 2006, the education minister taking over wrote, “with the [publication of the] PISA results, the scene was set for a national battle over knowledge in our schools…for those of us who had just taken over political power…the PISA results offered a ‘flying start.’ (Bergesen quoted in Sjøberg, 2014: 4). Similar to earlier, in 2006 the “Knowledge Promotion” reforms the minister implemented reflect closely the subject areas PISA tests.

Status shocks to status competition

This paper suggests the PISA rankings engender status concerns representing a peculiar type “external shock” (Kingdon, 1995). Kingdon theorizes that external shocks open a “window” for government actors to undertake large scale reform. One interesting thing about PISA, and rankings, is that the window it opens may stay open by virtue of the way they package PISA test knowledge: in rankings. As noted above, one can improve one’s performance in absolute
terms and still move down a ranking if the other competitors improve by more.

Notwithstanding an unprecedented improvement or a public rejection of the legitimacy of PISA, its rankings could provide the discursive resources to be used to legitimize education reforms for as long as Norway does not achieve top spot. Indeed, the PISA rankings seem to have induced in Norway a “race to the top” (Bieber and Martens, 2011: 103) whereby Norway engages in policies designed to compete in the OECD defined education game. Writing in 2014, Sjøberg indicates this may have occurred: reform of the Norwegian education system continues to use PISA scores as its rationale, and perhaps troublingly for Norwegian democracy, it continues to use PISA as its guide for reform (Sjøberg, 2014).

Conclusion

This article has theorized how IO rankings constitute a novel knowledge packaging technology that seeks to induce status competition in their target population. The Norwegian PISA case was used to explore this theorization because it satisfies so many of the criteria for status shocks and competition to emerge. The exploratory case study suggested overwhelmingly that PISA induced status shock and urgent policy measures designed to compete better against its rivals in the PISA rankings. However, most rankings organizations do not have the same degree of legitimacy, credibility, or publicity as PISA. Nonetheless, in the absence of material incentives attached to the growth of most of the new rankings, inducing status shocks and status competition seems to be one mechanism through which they could affect change, even if they do not (yet) work to the extent that PISA does. Thus, this paper has laid the groundwork for the investigation of 1) how or whether other theoretically “weaker” rankings can operate as status conduits and induce status competition 2) the process through which global rankings seek to improve their ability to induce status competition. Finally, 3) this paper indirectly points to further investigation into the ethical questions of when and why it might be preferable to package knowledge in transitive rankings, and when and why it might be preferable to package knowledge about populations as ratings based upon absolute thresholds of performance.

References


