

PISA, Policy and Persuasion: translating complex conditions into education "best practice"

Euan Auld* and Paul Morris

Department of Education, Practice and Society, UCL Institute of Education, London, UK

Abstract

Education reform is increasingly portrayed as a means to improve a nation's global competitiveness as measured by its performance in international league tables of pupil achievement. This has created a demand for comparative research which identifies 'what works' in high-performing school systems. A diverse array of consultancies, think tanks, and entrepreneurs has emerged to satisfy that demand, portraying their approach as a pragmatic and objective form of evidence-based policymaking. However, the attempt to translate complex conditions into straightforward solutions (i.e. 'what works') leads researchers into a basic paradox. This paper identifies the strategies used to address the core issues and to advocate reforms. We demonstrate that, though they are persuasive, the strategies fundamentally fail to overcome the problems inherent in the enterprise.

Introduction

The move towards evidence-based policymaking in recent decades has been accompanied by the emergence of new and influential actors in the field of Comparative Education. They include a diverse range of international organizations, think tanks, commercial consultancies, and policy entrepreneurs that use cross-national achievement tests (e.g. PISA; TIMSS; PIRLS) to construct and promote knowledge of transferrable 'best practices'. Whilst the task of identifying transferrable practices has been central to the field of Comparative Education since its beginnings, the contemporary approach differs significantly, especially in its reliance on cross-national achievement tests and its willingness to strongly advocate the transfer of specific policies and practices that 'work'. In so doing a basic paradox must be confronted, namely:

Whilst identifying policies for transfer relies on straightforward and generalisable causal claims that focus on school systems' practices and structures, the reasons underlying different levels of pupil achievement are inherently complex and explanations are conditional.

This paper identifies how that paradox is navigated through an analysis of the strategies used to develop knowledge claims and advocate policy actions. Whilst

* Corresponding author. Email eauld@outlook.com

scholars have highlighted the reductionist nature of the approach in specific publications (e.g. Alexander 2010; Coffield 2012), Demszky and Nassehi (2014) claim that we still have little insight into how ‘doing knowledge’ actually works (125). Cowen (2014) sharpens the problem, noting that though we know ‘the changing intellectual shape of the field... we are less good at seeing our fields of study in terms of its sensitivity to domestic and international political powers’ (285). This paper addresses that void. It builds upon our earlier analysis of the distinctive characteristics of contemporary policy-oriented comparisons, which we identified as a comparative genre (Auld and Morris 2014). Publications within the genre share a common communicative purpose, namely: *the identification of education ‘best practices’ (i.e. ‘what works’) to support policy transfer.*

Individuals and organisations engaging with this communicative purpose form a discourse community, responding to and reinforcing a new grand narrative which emphasises the role of education in preparing nations to compete in the global knowledge economy. The story logic follows an economic rationale and is identified as follows: (1) education’s primary function is to prepare individuals/nations to compete in the global knowledge economy; (2) outcomes measured by international surveys (e.g. PISA) provide a reliable proxy for the quality of a system’s human capital; (3) reform must therefore focus on improving education outcomes relative to international competitors (*the development of world class schools*); (4) this requires the identification (and transfer) of ‘what works’/‘best practices’ in high-performing systems.

Although the search for transferrable policy solutions in the form of ‘what works’ draws legitimacy from positivistic science’s claims to identify ‘true causes’ (Lemke 1995), education research generally deals with issues that are ‘trans-scientific’ (Weinberg 1974). That is, they have an inherent moral component, studying them is too expensive or impractical to replicate, or, our current focus, the complex nature of the subject matter precludes the identification of clear (and generalizable) lines between cause and effect. As Fischer (1993) notes, ‘any particular set of facts could- at least arguably- be consistent with a variety of theories’ (32) (as the many features proposed to explain the high performance of East Asian nations illustrates), and the identification of ‘best practices’ requires the privileging of one of innumerable possible ‘causal stories’ (Stone 1989).

This evokes the distinction between research and advocacy (Majone 1989). While the former tends to develop complex and conditional explanations, the latter seeks to identify straightforward solutions (i.e. ‘what works’) that can be used to influence policy. Although this presents policymakers with a direct course of action, the translation of complex conditions into simple solutions leads advocates to commit ‘the fallacy of the simple question’ (Goldstein 1993), and advocating policy transfer relies on a misapplication of syllogistic reasoning. Biesta (2010) points out that ‘much talk about “what works” is premised on the assumption of closed, deterministic systems... and a mechanistic ontology’ (497). The resulting conflict cannot be

resolved by way of methodological innovation. Consequently, alternate strategies are required to navigate the paradox and to promote reforms.

Given the move from research to advocacy, we turned to political science to aid interpretation of the strategies, focusing on the role of narratives in persuasion. Stone (1988) identifies two broad storylines used in policy texts. First, the ‘story of decline’ (or, the need to improve) provides impetus for reform. Second, the ‘story of control’ inspires faith that the identified problems can be fixed, providing policymakers with authority for action in the form of expert knowledge of ‘what works’. While both elements are used to develop persuasive texts, it is the story of control that invites the identified paradox. Preliminary analysis revealed that though researchers indicated awareness of the issues associated with transfer, they duly attempted to suppress or marginalize them to enable delivery of the communicative purpose (Auld and Morris 2014, 140-147).

Seeking to understand this contradiction, we turned to the advocacy coalition framework’s (ACF) (Sabatier and Jenkins-Smith 1993) tripartite hierarchy of beliefs: deep core, policy core, and secondary beliefs. These beliefs are implicit in the narrative used to promote reforms. The deep core is positioned as the most fundamental and least flexible, and is defined by Jenkins-Smith, Silva et al. (2014) as the ‘normative and ontological axioms that shape an individual’s beliefs’ (485). The policy core is the focus of this paper and is defined as the ‘relatively abstract beliefs concerning the underlying causes of a problem within a subsystem, or general strategies for dealing with a class of problems’ (498). Secondary beliefs are represented in advocates’ specific policy preferences (486) and therefore hold the greatest potential for variation within the community. The relevant beliefs can be summarised as:

- Deep core: *closed system and mechanistic ontology*
- Policy core: *generalizable causal relationships, located within school systems’ practices and structures*
- Secondary: *specific policy preferences (variable)*

Authors of the featured publications do not necessarily collaborate, and often promote conflicting policy solutions. In this respect they do not constitute an advocacy coalition. However, Jenkins-Smith, Nohrstedt et al. (2014) emphasise that ‘analysts [are] increasingly specializing in one or more of the [ACF’s] subcomponents’ (187), and McBeth, Shanahan et al. (2007) argue that the analysis of beliefs (as defined by ACF) and political strategies is ‘an unexplored area in which the two fields intersect and strengthen each other’ (90). ACF views actors as boundedly rational and susceptible to identity-protective cognition, perceptual filtering and bias assimilation. Sabatier (1998) clarifies the implications: ‘members will resist information suggesting their deep core or policy core beliefs may be invalid and/or unattainable, and usually

will use formal policy analyses to buttress and elaborate those beliefs (or attack their opponent's views)' (104-105).

We demonstrate how this process of buttressing, elaboration and selective interpretation is used to lend authority to the narrative of control. The analysis provides insight into the cognitive structuring of reports, revealing the habitual dissonance that emerges as actors attempt to develop straightforward causal narratives that protect their core beliefs. We argue that these moves and strategies are not unique to specific reports, but follow a common repertoire, and are necessitated by the nature of the research ambition. Moreover, by identifying the cumulative effect of these strategies, and relating it to the broader context, the paper provides insight into how "comparative education" is changing shape in response to global reform agendas.

Strategies for navigating complexity

We analysed texts that were representative of the genre (i.e. share the same communicative purpose), identifying the different moves and strategies used to deal with the above paradox. Sample reports were analyzed using principles from *move analysis* (Swales 1990), a subset of genre analysis, which 'explicitly studies texts in terms of their rhetorical goals and how they work to achieve those goals' (Tardy 2011, 56). This focused analysis on how advocates attempted to circumvent the core paradox, but also looked more broadly at the moves used to develop a persuasive narrative. ACF's concept of buttressing then provided a basis for interpreting these strategies.

The first stage of analysis focused on four key reports (Auld and Morris 2014), and these preliminary strategies were then used to guide the analysis of a broader corpus of texts. This continued in an iterative process, developing and nuancing understanding of the strategies until they began to reach saturation. A diverse and illustrative sample of reports was then selected for the presentation of the analysis in this paper (see Table 1.), focusing on texts which were: published by high-profile actors; influential (i.e. cited extensively by other advocates and/or by policymakers); and were targeted at both global and national audiences.

Fairclough (2003) argues that 'analysis of generic structure is of value for more strategic, purpose-driven texts', highlighting a 'need to look for staging in analyzing texts and interactions' while warning that we shouldn't expect 'to always find that they are organized in terms of a clear generic structure' (74). The analysis revealed a broad repertoire of strategies. These strategies were organized into the five core moves, shown in Table 2. The first two moves are recognized in the literature as important aspects of persuasive texts, and set the foundations for advocating knowledge claims. The final three moves, on which we focus, detail the strategies used to overcome the issues inherent in the policy core directly.

The moves and strategies were utilized to varying degrees and in varying combinations across the reports. Rather than a rigid template, the framework provides

a basis for engaging critically with the growing and influential range of ‘best practice’ reports. Below we illustrate and analyse each of the strategies.

Table 1. Main Publications Featured[‡]

Author/Organisation	Role/Type	Title/Year	Abbreviation
OECD	International organisation, committed to the promotion of policies that improve economic and social well-being	2012: <i>Lessons from PISA for Japan, Strong Performers and Successful Reformers in Education</i>	OECD (2012)
		2013a: <i>Lessons from PISA 2012 for the United States, Strong Performers and Successful Reformers in Education</i>	OECD (2013a)
		2013b: <i>PISA 2012 Results: What Makes Schools Successful? Resources, Policies and Practices (Volume IV)</i>	OECD (2013b)
Andreas Schleicher	Director for the Directorate of Education and Skills (OECD)	2009: "Securing Quality and Equity in Education: Lessons from PISA." <i>Prospects</i> 39: 251-263	Schleicher (2009)
		2012: <i>Preparing Teachers and Developing School Leaders for the 21st Century: Lessons from around the World</i>	Schleicher (2012)
		2014: <i>Equity, Excellence and Inclusiveness in Education: Policy Lessons from around the World</i>	Schleicher (2014)
McKinsey & Company	Global management consultancy	2007: <i>How the World's Best-Performing School Systems Come Out on Top</i>	McKinsey (2007)
		2010: <i>How the World's Most Improved School Systems Keep Getting Better</i>	McKinsey (2010a)
		2010: <i>Capturing the Leadership Premium: How the world's Top School Systems are Building Leadership Capacity for the Future</i>	McKinsey (2010b)
Sir Michael Barber [†]	Chief Education Advisor, Pearson	2009: <i>Impossible and Necessary: Are you Ready for This?</i> London, NESTA.	Barber (2009)
Institute for Public Policy	Think tank (UK)	2012: <i>Oceans of Innovation: the Atlantic, the Pacific, Global Leadership and the Future of Education</i>	IPPR (2012)
Grattan Institute	Think tank (AUS)	2012: <i>Catching up: Learning from the Best School Systems in East Asia</i>	Grattan Institute (2012)
Global Education Management Systems	Global education consultancy	<i>The Efficiency Index: Which Education Systems Deliver the Best Value for Money?</i>	GEMS (2014)
Marc Tucker	President, National Center on Education and the Economy (US)	2011: <i>Standing on the Shoulders of Giants: An American Agenda for Education Reform.</i>	Tucker (2011)
D. Reynolds and S. Farrell	Commissioned by Ofsted	1996: <i>Worlds Apart: A Review of International Surveys of Educational Achievement involving England</i>	Worlds Apart (1996)

[†] In addition, Barber co-authored each of the featured McKinsey Reports, and co-authored IPPR (2012) with Pearson colleagues, Katelyn Donnelly and Saad Rizvi.

[‡] The publications listed here are not duplicated in the references, and are cited throughout the paper according to the stated abbreviations.

Table 2. Overview of Moves/Strategies

Move/Strategy	Overview of function
<p>1. Establishing the frame: <i>the logic of action</i> a) constructed paranoia & crisis rhetoric b) appeal to precedent c) undermine critics/alternate perspectives</p>	<p><i>The context is framed to present the rationale for reform and the value of the research ambition as self-evident. Further strategies stress its significance/urgency and undermine alternate perspectives. Accepting the frame (and, therefore, the research ambition) is portrayed as the only reasonable position, and the reader is confronted with the logic of action.</i></p>
<p>2. Establishing expertise a) reputation and/or experience b) statistics and quasi science c) genre knowledge (academic literature)</p>	<p><i>Scholarly/scientific language is used and awareness of academic methods/conventions is demonstrated, providing an authoritative platform for advocating reforms. The advocate's reputation and/or experience is emphasized and endorsements from high-profile actors are provided.</i></p>
<p>3. Restricting the analytical focus a) focus on quantifiable outcomes b) focus on high-performing systems c) confirmatory cases d) selective interpretation e) incomplete stories f) focus on overarching policy levers g) focus on factors 'amenable to control'</p>	<p><i>Cases and/or lines of enquiry that investigate complex causal interactions are marginalized or excluded, and explanations privilege factors that are amenable to control. This lays the foundations for the development of straightforward (and generalizable) causal stories.</i></p>
<p>4. Drawing Recommendations Step one (highlight issues) Discuss limitations; include caveats</p> <p>Step two (deal with issues) a) "established knowledge" (outsourcing)</p> <p>b) reject limitations/caveats</p> <p>c) circumvent limitations/caveats i) hanging observations ii) softening the ambition iii) obfuscation</p>	<p><i>By demonstrating awareness of the issues/limitations, critique is pre-empted and the credibility of the researcher is safeguarded.</i></p> <p><i>Causal claims are attributed to other sources, or presented as 'what we already know'. In this way, responsibility for dealing with limitations can be outsourced.</i></p> <p><i>The limitations are rejected or restated as of secondary importance, enabling the research ambition to be delivered directly but leaving open the charge of over-simplification.</i></p> <p><i>Commonalities/associations are noted but explicit causal claims are avoided. The research ambition is thus delivered indirectly, protecting the advocate from criticism.</i></p>
<p>5. Qualifying recommendations a) necessary (but not sufficient) conditions b) knowledge is contextual c) the burden of implementation d) knowledge is imperfect</p>	<p><i>Complications and caveats are revisited or introduced to finesse the knowledge claims and/or to deflect responsibility if a policy based on the advocated 'best practice' fails.</i></p>

Move 1. Establishing the frame: *the logic of action*

The process of strategic framing is often identified as the key stage in the policy process (Sabatier 1998), determining what conditions may be considered as problems that require attention, privileging specific values and objectives, and setting the story on the desired trajectory. The frame logic lays the foundations for the reform narrative, situating education in the context of unrelenting global economic competition. Reports commonly employ three strategies (*crisis rhetoric; appeals to precedent; and, undermining critics/alternate perspectives*) to establish and reinforce the frame, demonstrating the necessity and viability of the research ambition. These strategies combine to form *the logic of action*. Though they feature prominently at the beginning of reports, they are often revisited to provide impetus for reform and to support specific recommendations.

The role of crises in initiating policy change is emphasized in political science (e.g. Stone 2012). Two distinct levels of crisis development operate. The first creates the conditions for a crisis, a move which Oates (2013) describes as ‘constructed paranoia’. Second, reports targeting a specific context highlight the system’s performance relative to its competitors to establish a crisis of standards. Kallo (2009) highlights the large number of publications that the OECD releases on ‘globalisation, economics of education, and human and social capital’ (182), and we focus on how the OECD markets the above logic, effectively engaging in a form of ‘frame sponsorship’ (Rein and Schon 1993). In *Lessons from PISA for Japan* (2012), the OECD asserts that ‘values and preferences evolve and education systems must change to accommodate them’ (23). It presents the reasons for this evolution:

Rapid globalisation and modernisation are posing new and demanding challenges to individuals and societies alike... High-wage countries will find that they can only maintain their relative wage levels if they can develop a high proportion of such knowledge workers and keep them in their work force... This is not a description of one possible future, but of economic dynamics that are now in play... The implication is that the yardstick for educational success is no longer simply improvement against national standards, but against the best-performing education systems worldwide. (16)

This new yardstick introduces the concept of a universal standard of quality. As Kamens and McNeely (2010) note, ‘in a world where national educational systems are viewed as unique in structure, history, and purpose, international testing would have little plausibility’ (8). Education is defined entirely in economic terms, with regard to the instrumental role that it plays in developing human capital, and empowering knowledge workers. Building on this image, Schleicher (2014) explains the significance of PISA outcomes:

According to one estimate, if all 15-year-olds in the OECD area attained at least level 2 in the PISA mathematics assessment, they would contribute over USD 200 trillion in additional economic output over their working lives. (21)

The source of this estimate is another OECD report (2010a). By positioning PISA as a reliable indicator of an individual’s ability to ‘meet real life challenges’, and as a

proxy for a nation's stock of human capital, education systems are thus rendered directly commensurable, and PISA is established as the 'world's premier yardstick' for education quality (OECD 2013, 11). Gorur (2015) refers to this process as 'producing calculable worlds', evoking Stone's (1988) claim that 'much... counting is an effort to identify a statistical community in order to demonstrate common interests and thereby stimulate creation of a natural community' (135). This community coalesces around a cycle of cutthroat competition. Henry, Lingard et al. (2001) claim that the OECD is primarily interested in 'fostering rather than mediating market relations'. This interest in *fostering* market relations is particularly evident in the evangelism of Andreas Schleicher. In his foreword to the McKinsey (2007) report, Schleicher explains:

The capacity of countries - both the world's most advanced economies as well as those experiencing rapid development - to compete in the global knowledge economy increasingly depends on whether they can meet the growing demand for high-level skills. This, in turn, hinges on significant improvements in the quality of schooling outcomes...

The foreword concludes with a stern warning:

The world is indifferent to tradition and past reputations, unforgiving of frailty and ignorant of custom or practice. Success will go to those individuals and countries which are swift to adapt, slow to complain and open to change. (6)

A subsequent challenge is amended according to the audience:

McKinsey Report, Foreword (global): the task for *governments* will be to ensure that *countries* rise to this challenge (2007, 6)

Lisbon Council (Europe): the task of *European* governments will be to ensure that *European* countries rise to this challenge (Schleicher 2006, 16)

UK[§]: the task of *UK* policymakers is to help *its citizens* rise to this challenge.

These incessant challenges capture the spirit of 'constructed paranoia'. In the context of this global war for talent, education serves as nations' munitions factory and 'improvement' becomes a matter of national/system survival. Reform is the policymaker's patriotic duty. Reports targeted at a specific context either highlight a system's poor performance relative to its competitors, developing a crisis of standards, or emphasize that 'every country has room for improvement, even the top performers' (OECD 2013a, 13; 2013b, 4)

Michael Gove, former Secretary of State for Education in England, illustrates how the above conditions are sharpened into crisis to provide impetus for reform:

We are still falling further behind the best-performing school systems in the world... leaving our children behind in the global race... That matters because business is more

[§] Schleicher A., 16.11. 2012, Times Educational Supplement, You must emulate and innovate to keep pace: <https://www.tes.com/article.aspx?storycode=6301764>

mobile than ever, and employers are more determined than ever to seek out the best-qualified workers. Global economic pressures... are driving all nations to pursue educational excellence more energetically than ever before. And today's league tables show that nations which have had the courage radically to reform their education systems - like Germany and Poland - have significantly improved their performance. (Gove 2013)

This leads to the next strategy, *appeals to precedent*. The four main sources of precedent used are: *other (high-performing) systems*; *high-profile actors/texts*; *other fields/disciplines*; and, *a bygone era*. Each of these reinforces the frame, promoting the necessity and viability of the research ambition as part of the logic of action.

References to *other (high-performing) systems* operate as a key stage in constructed paranoia, challenging policymakers to ignore their global competitors and risk being left behind. Having highlighted the urgency of the situation, the OECD (2013a; 2013b) explains:

More and more countries are looking beyond their own borders for evidence of the most successful and efficient policies and practices. (12; 3)

And Tucker (2011):

Every one of the top performers is very conscious of what the other top performers are doing, though some benchmark more aggressively than others. (5)

Second, references to *other influential or high profile actors/texts* serve to legitimize the ambition. For example, Schleicher (2009) asserts:

Several expert studies have reviewed *some* of the education systems with high performance standards and found *some features they share*. (254 italics added)

The expert studies referred to are published by the OECD (2004) and McKinsey (2007), the latter of which relied heavily on the former to support its claims. Similarly, the Grattan Institute (2012) asserts:

A body of international research has identified the common characteristics of high-performing education systems. (13)

Again, reports by the OECD and McKinsey are referenced to support the claim. None of these reports deals with the underlying issues, especially establishing causality. Silova and Brehm (2015) note that such a process of self-referencing, 'although not problematic in itself, can create a closed system of circular knowledge - the legitimation of a view by claiming another scholar has published the same view elsewhere' (27). We return to examine the implications below (see *establishing causality*).

Third, references to *other sectors (e.g. business) or fields (e.g. medical research)* are used to promote and justify the research ambition:

Variations in cultural context and tradition have never prevented management in any

area from trying out reforms that have been introduced abroad. (Worlds Apart 1996, 59)

Along the same lines as “evidence-based medicine”, Long Beach Unified School District identifies the best delivery methods from pilot data and then rolls out the program across all its primary schools. (McKinsey 2010a, 42)

Such appeals boldly dismiss the possibility that different objects require different methods. A final variation is provided by Tucker (2011), who employs what Stone (1988) terms an appeal to a *bygone era*:

A century ago, the US was among the most eager benchmarkers in the world. We took the best ideas in steelmaking, industrial chemicals and many other fields... At the same time, we were borrowing the best ideas in education... It was the period of the most rapid growth our economy had ever seen...

But, after World War II... we evidently came to the conclusion that we had little to learn from anyone. As the years went by, one by one, country after country caught up to and then surpassed us in several industries and more or less across the board in precollege education. And still we slept. (1)

Those not engaged in the crusade are implicitly portrayed as backward, arrogant or complacent. We now turn to explore the more direct forms of this strategy of derision. Questioning and parodying the *logic* of critical or alternate perspectives, or questioning the *motives* and/or *expertise* of critics deflects attention from the issues inherent in the search for best practices.

Advocates of ‘what works’ often depict themselves as ‘pragmatists’, thereby drawing implicit contrast with “impractical” theorists (Farnsworth and Solomon 2013). Any opposition to ‘evidence-based’ interventions is thereby portrayed as irrational (Hammersley 2013). Advocates commonly attempt to undermine the logic of alternative viewpoints:

The situation in which England finds itself is now so worrying, that the risk involved in looking outward and trying new practices is worth taking... *The way to cease being worlds apart is surely to adopt an open mind.* (Reynolds and Farrell 1996, 59 italics added)

The capacity of educators to stumble into a false dichotomy and debate it (vigorously and at length to the benefit of no-one) is legendary – for example, the widely held but absurd view that because some things can’t be measured, we should measure nothing. (Barber 2009, 19)

While the former of these statements portrays any doubt as narrow-minded dogmatism, the latter derives legitimacy by attacking a straw man, for few would argue that we should measure *nothing*. By distancing themselves from absolute transfer (see *genre knowledge*) and deriding this extreme negative position, advocates are able to adopt the middle ground within the established frame; the only reasonable position.

Hopmann and Brinek (2007) relate how their *motives and expertise* were

questioned by the German PISA consortium, which declined an invitation to contribute to a volume engaging critically with PISA data because it would: ‘provide a forum for unproven allegations’; critics ‘were unqualified to discuss PISA’; and, ‘they were probably driven by envy or other non-scholarly motives’ (14). Barber (2009) provides an example, questioning the motives of critics:

There are many educators and system leaders who simply don’t believe that successful change is possible. There are academics who use sophisticated statistical techniques to support the view that social background remorselessly determines outcomes, regardless of what education systems do. (19)

Here generalised assertions of unclear provenance are used to undermine the *ethos* of anyone who emphasizes (probably using PISA data) social background as an influence on student outcomes. This strategy reflects ACF’s emphasis on the importance of shared beliefs for in-group cohesion, and distrust of others who ‘since they come to conclusions different from ours must be motivated by hidden, nefarious interests’ (Sabatier 1998, 109).

These strategies combine to form the *logic of action*: the need for urgent reform is established; appeals to precedent demonstrate both the significance and the viability of the research ambition; and, those who question the ambition are irrational dogmatists, lack expertise, or have a hidden (nefarious) agenda.

Move 2. Establishing expertise

Perceptions of advocates’ credibility is a vital aspect of effective persuasion (Oxley, Vedlitz, and Wood 2014). This second move is pursued indirectly by undermining critics, and below we highlight three strategies used to establish expertise (*reputation and/or experience*; *scientific credentials*; and *genre knowledge*). Credibility is also affected by whether the message resonates with individuals’ pre-existing values and beliefs (Lachapelle et al. 2014), underlining the limitations of advocate’s influence in contexts where its core values are not shared, and the primary importance of establishing the frame. Having demonstrated the necessity and viability of the research ambition, advocates must still convince the reader to privilege their explanations over those of other experts.

Reputation, past success and/or experience, and *insider knowledge* rather than demonstrable rigour, are used to support claims and to undermine alternate perspectives. The OECD brand is used to justify policy interventions, while Andreas Schleicher has been hailed as both ‘the world’s schoolmaster’ (Stewart 2011), and ‘the godfather of... global education comparisons’ (Coughlan 2012), with his proclamations often quoted reverentially by policymakers, the media and academics. Entrepreneurial academics such as Michael Barber develop reputations as celebrity reformers, moving across nations to promote their improvement formulas. Indeed, Michael Gove (2011) lauded Andreas Schleicher and Michael Barber as two of the most important men in world education.

Individual reports may emphasize the credentials of the authors and/or organisation, for example:

This paper is the result of constant dialogue among the authors as we've worked together, first on education reform in Pakistan... and second as part of an innovative team at the heart of IPPR, the world's largest education company, where we are seeking to resolve the dilemmas of providing quality education to people of all ages on every continent. (IPPR 2012, 1)

This report is the result of a collaborative effort between the OECD and international experts with extensive expertise in analysing the performance of education systems internationally (OECD 2012, 5)

Alternatively, reports feature endorsements from other members of the community, or acknowledge high-profile actors as advisors. For example, Schleicher provides a foreword for both The McKinsey (2007) report and GEMS (2014), while Grattan Institute (2012) states that its research was based on a roundtable discussion (7), providing details of the high-profile attendees (including Schleicher and representatives from high-performing East Asian systems) whose insights were used to 'direct' the analysis. Ball and Exley (2011) note how validation from 'academics within prestigious universities lends legitimacy, authority and an air of rigour to ideas' (161). The McKinsey (2010a) report, for example, includes a glowing foreword by Professor Michael Fullan, who is acknowledged for his 'counsel and thought partnership'(3).

Notwithstanding reputation and experience, the demand for evidence requires advocates to demonstrate their *scientific credentials*. The OECD draws heavily on its reputation for statistical expertise, stating:

[PISA] provides the world's most extensive and rigorous set of international surveys of the knowledge and skills of secondary school students. (OECD 2013a, 3)

Schleicher (2014) repeatedly refers to 'what PISA shows', using associations derived from the data to promote policy ideas (see *hanging observations*), and often concludes presentations by stating: 'Remember, without data you are just another person with an opinion'. Waldow (2012) describes this as 'a classic example of externalisation to the principles and results of science, as a way of legitimizing oneself and undermining one's critics' legitimacy' (423).

Barber creates a 'conceptual formula' for developing well-educated students (Barber 2009; Barber and Mourshed 2009), which the IPPR report (2012) promotes as a platform for curriculum reform rather than a straightjacket (52):

To clarify a debate about the curriculum, which has a tendency to spiral into jargon, we have attempted to summarise what children should learn in a simple mathematical equation: Well-educated [students] = E(K+T+L). (IPPR 2012, 49)

Where K equals *knowledge*; T, *thinking/thought*; L, *leadership*; and E, *ethics*. The "science" of *deliverology* is promoted by McKinsey (see Barber, Kihn et al. 2011), a

term coined by Barber, while IPPR (2012) asserts: ‘increasingly, a science or quasi science of effective delivery in government is emerging’ (60). This quasi-science underpins the concept of ‘system effectiveness’ that is being applied by ‘increasing numbers of system leaders’ (Barber 2009, 16). It also ties in well with McKinsey & Company’s unequivocal motto, “everything can be measured, and what gets measured gets managed”.

Despite the primacy of quasi-scientific method, reports are mindful to stress their qualitative aspects. For example, the OECD (2012) highlights that the research entailed a process of document review and then interviews, which included historians, economists, education experts, citizens, and journalists, thereby allowing for an ‘alternative benchmarking’ (23). Similarly, the McKinsey (2010a) report draws attention to its interviews to emphasize that ‘findings [were] not ... the result of some abstract statistical exercise’ (13).

The association with crude processes of selective policy ‘borrowing/copying’ remains problematic. By directly denouncing such practices, awareness of the relevant academic literature (*i.e. genre knowledge*) is also demonstrated. For example, in the Grattan Institute (2012) states:

The report does not claim that the political and policymaking structures of East Asia can or should be reproduced elsewhere. (2)

Reports now generally describe a process of policy *learning*, though the actual approach is rarely articulated. For example, Schleicher, (2012a) describes the reform process in Shanghai:

The idea was not to *copy* what they were doing, but to *learn* from them... Though one can always question whether policies that are successful in one place will succeed in another - and surely no country can simply adopt another nation’s system or policies - comparative data and analysis seem to rapidly expand the scope for *learning* from the successes and failures of education policies and practices around the world. (81, italics added)

Learning is interpreted differently across reports, and it is necessary to determine whether the shift is genuine, strategic, or a misapplication of the term. Genre knowledge is further demonstrated by highlighting limitations and including caveats relating to culture, context, the complexity of social interactions, the problem of causality, and even direct statements warning against policy transfer. For example:

It is important to recognize that there are contextual differences between systems, and that what works in one context may not work in another. (McKinsey 2010, 3)

By itself, a cross-national international assessment such as PISA cannot identify clear-cut cause-and-effect relationships. (Schleicher 2009, 253-254)

There is no single combination of policies and practices that will work for everyone, everywhere. (OECD 2013a, 1; 2013b, 4)

Whilst identifying limitations and including caveats provides some protection from critique, it also highlights the problems inherent in the research ambition. The challenge therefore lies in demonstrating genre knowledge while developing conclusions that can be packaged into direct policy recommendations (see *Move two: deal with issues* below).

These strategies contribute to the overall persuasiveness of a text, and lay the foundations for advocating the transfer of identified best practices. Although successive analyses have demonstrated the selective referencing of this evidence in England (Morris 2012; Yun and Morris 2015), elsewhere advocates attempt to forge new markets, both retailing education solutions and offering the prescribed services (Ball 2012). Next we examine strategies used to deal directly with the core issue; namely, *that whilst data on pupil performance may 'reveal' possible correlations, it provides no direct basis for establishing the causal connections required to isolate and transfer policy interventions.*

Move 3. Restricting the analytical focus

By restricting the focus to the *quantifiable outcomes* measured by international surveys, broader outcomes and the effects of specific policies, including unintended consequences, are excluded. Rappleye (2010) describes how complexity is collapsed in this way (along economic lines) in the context of education and development and how this both frames the way in which the process is conceptualized (marginalizing broader effects of educational interventions), and how causality is then hypothesized ('judgment calls') in the pursuit of progress. Barber's (in Ozga 2014) account of his time with the Prime Minister's Delivery Unit (PMDU) explains how inquiry is narrowed:

Because we had some targets or goals that were public, we started from the targets. So we worked back from a target... we didn't go into it with a kind of open research point of view where we say what are the many questions we could ask about this data... we're going in with a particular perspective. (21)

GEMS (2014) also acknowledges its selectivity:

We have focused on one output (system efficiency), but the choice of this output in each context is an important part of the debate. (30)

Notwithstanding this *preliminary narrowing (focus on quantifiable outcomes)*, the features of complex and diverse systems must be harnessed to provide a straightforward causal story that is consistent with the community's core beliefs. The following strategies (*focus on high-performing systems; confirmatory cases; selective interpretation and incomplete stories; overarching policy levers; and, factors that are amenable to control*) narrow the evidential base to *cases* that confirm the causal story and to *features* that are amenable to direct intervention.

Reports may focus on any number of "*high-performing systems*" (e.g. cities, nations, provinces or regions), whereas poor performers are excluded. This streamlining

facilitates a straightforward narrative of success. The focus is often revealed in the title of reports, (e.g. McKinsey 2010a; Grattan Institute 2012), and is clearly stated, for example:

The review represents one of the first attempts to compare school leadership across a range of high-performing education systems. (McKinsey 2010b, 3)

The subsequent chapters present detailed analyses of high-performing education systems - Canada (Ontario), Finland, Singapore, and China (Hong Kong and Shanghai) (OECD 2012, 17)

Greene (2012) notes the basic methodological flaw: *selection on the dependent variable*, and explains, ‘when you look at the things that successful organizations are doing, you have no idea whether each one of those things caused the good outcomes, had no effect on success, or was actually an impediment that held organizations back from being even more successful... by avoiding variation in the dependent variable it prevents any scientific identification of causation’.

The narrative of success is broadened by including ‘most improved systems’ (McKinsey 2010a), or ‘successful reformers’ (OECD 2012; 2013a). The preferred policy intervention is then associated with any improvement in outcomes (success story), regardless of overall performance. Schleicher (2014) goes further, focussing on systems that ‘*do most to achieve excellence, equity and inclusion*’ (11 italics added). Demonstrable improvement is no longer necessary, and the preferred policies are promoted through a narrative of good intentions.

By delimiting the analysis of high performing systems to *confirmatory cases*, commonalities can be presented unconditionally as ‘best practices’, or ‘what works’. Likewise, poor performers that share the identified ‘best practice’ may be acknowledged but are subsequently dismissed. Alexander (2012) highlights the omission of deviant cases from a number of influential reports, and the selectivity is confirmed in McKinsey (2010b):

The education systems in these countries all perform strongly on international tests, or their performance in this respect is improving, and *they demonstrate good practices in school leadership*. (4 italics added)

The circularity (they were selected because they demonstrate what is considered good practice) is confirmed:

The international survey conducted as part of this research largely *reconfirms* existing knowledge about the roles which effective school leaders play. (7 italics added)

This process of ‘selective exposure’ (Jones, McBeth and Shanahan 2014, 12) evokes the important distinction between how conclusions are *developed* and *justified* (Majone 1989). Though causal claims may be developed by other means, the symbolic power derived from reference to high-performing systems encourages their use as the *medium* of persuasion, a strategy that encourages confirmatory referencing.

The narrative of success (what high-performing systems *do*) is undermined by any divergent case. This is particularly evident in exploratory reports, in which multiple lessons are derived from a range of cases. *Worlds Apart* (1996) directly confronts the issue. Having observed that ‘an assertive principal’ does not have the expected association with school effectiveness in the Netherlands, it asks:

What is it in the local, regional or national ecology that might explain this finding?... Answering this question inevitably involves the generation of more complex explanations that operate on a number of levels, and is likely to generate more complex theoretical explanations. (5)

The question is duly shelved. Such questions are pre-empted by acknowledging and subsequently excluding divergent cases, and then reintroducing them when they support the narrative. For example, Tucker (2011):

We know that the complete transformation of the whole system of policy and practice we have suggested will seem an overwhelming prospect to many people. So we turn to Canada as our best example of a country that might be used as a source of strategies for making great improvements in the short term. (2)

Canada is subsequently dismissed as an outlier because it does not share the same ‘system of gateways’ and curriculum arrangements as ‘virtually all’ the cases. Later, however, it is rehabilitated to support Tucker’s *Agenda for American Education*:

We have not mentioned Canada much until now, because this is where it fits. (43)

The strategy reflects ACF’s emphasis on ‘biased assimilation’, whereby individuals ‘tend to interpret evidence in a way that supports their prior beliefs and values’ (Henry 2011, 365). Alternatively, by avoiding divergent cases altogether the focus is restricted to cases that share the selected feature(s). An *incomplete story* is constructed, in which the feature/policy is introduced and then endorsed by a vague number of cases (e.g. ‘some’, ‘most’, ‘several’ ‘nearly all’), or by providing examples from specific cases while excluding others:

In the pages that follow, we will point out when *all* appear to share a policy framework, when *most* do and when *some* do. (Tucker 2011, 2)

The elements vary across the education systems described, but *generally include...* (OECD 2012, 17)

Selecting and emphasizing features and trends in this way tackles the problem of diversity, orchestrating and overstating agreement, and allowing incomplete stories to be combined to form a coherent master narrative. These incomplete stories function as *hanging observations* (see below).

While exploratory studies (e.g. McKinsey 2007; 2010a; Tucker 2011) develop an overview of high-performing systems and then accentuate common features, targeted studies restrict inquiry to an *overarching policy lever*, which is defined by the Grattan Institute (2012) as: ‘an element or characteristic of a system that can be changed in

order to achieve a strategic objective' (29). Common levers are leadership (e.g. McKinsey 2010b), curriculum (e.g. DfE 2012), and teacher quality (e.g. Schleicher 2012). Other reports focus on a number of policy levers but deal with each in turn (e.g. Schleicher 2014). By privileging one overarching policy lever, the significance of other variables may be acknowledged but then dismissed, or dealt with separately. Alternatively, other variables may be ignored, but with the concession that the selected policy lever cannot in itself guarantee improved outcomes (see *necessary but not sufficient conditions*).

Analyses which focus on a single overarching lever are better positioned than exploratory studies to select cases that confirm the preferred story. When confronted by variation between cases on the selected policy lever, two main strategies emerge. First, and as above, they emphasize aspects of these cases that confirm the story while overlooking anomalous characteristics (e.g. by focusing on historical rather than contemporary arrangements in some systems, but not others). Alternatively, the variety of policies in different systems are collated under the same authoritative banner, highlighting 'key' approaches in selected cases to develop sub-stories. The basic claim is that, despite some variation, there remains broad consensus in high-performing systems about the significance of the overarching policy lever(s).

Diversity within an overarching policy lever may be presented positively, as illustrating the options available to policymakers. Schleicher (2012a, 81) notes the 'major unresolved issues on effective teacher policies, both within and between countries', but asserts that 'Summit participants agreed that significant improvement is possible' and, 'the many examples of reforms in this publication... show that challenges can be successfully addressed' (precedent in *other systems*). The impact of the specific policy actions in these contexts, and how they operate amidst the broader array of social mechanisms, is unexplored. Complexity is subsumed under the authority of the overarching policy lever.

In terms of the actual content of explanations, the promotion of 'what works' requires advocates to focus on the features of education systems that are *amenable to control*, and to underplay broader influences on student outcomes (e.g. culture, history, or tuition). The Grattan Institute (2012) illustrates:

Success in high-performing education systems in East Asia is not always the result of spending more money... Nor is success culturally determined, a product of Confucianism, rote learning or Tiger Mothers. (2)

Neither cultural difference nor Confucian values can explain how, in just five years, Hong Kong moved from 17th to 2nd in PIRLS. (12)

Success is also not driven by the size of the system... the four high-performing education systems in East Asia vary in size. (12)

High performance in education systems in East Asia comes from effective education strategies that focus on implementation and well-designed programs that continuously improve learning and teaching. (12)

The improvement in outcomes is used to dismiss the significance of conditions outside schools (even as *enabling factors*), with *any* variation between the cases on certain variables (e.g. size) used to infer that they have had *no influence* on outcomes. In contrast, differences in the ‘education strategies’ across contexts are accommodated within the narrative as options, while the evidence that East Asian heritage students perform well even when educated outside these high-performing systems is overlooked (see Feniger and Lefstein 2014).

Finally, contextual elements such as family background are marginalized by distinguishing them from features that are amenable to control, which are described as ‘operational elements’ (OECD 2012, 23). These operational elements are further disaggregated into constituent (but related) parts. The distinction compartmentalizes complex systems into discrete ‘factors’, legitimizing the strategies (above) used to restrict the analytic focus, and acknowledging but then subordinating contextual influences. This process of buttressing implies a reversal of philosophical inquiry, allowing advocates to (re)interpret the nature of social reality in a way that supports the research ambition.

Move 4. Drawing recommendations: making causal inferences

Having narrowed the focus onto the selected variable(s), reports move to develop policy recommendations. Unless advocates can establish a causal claim the foundations of the enterprise are undermined. Commonly, a paradoxical pattern of moves unfolds as advocates simultaneously eschew direct causal claims whilst identifying policies for transfer.

Step one: highlight issues

The first step anticipates criticism by highlighting limitations and providing caveats (see *establish expertise* above). Few discuss the problems of establishing causal relationships explicitly, or the complexity of social systems. Unlike the issues of culture and context, they cannot easily be recast as secondary considerations, and once acknowledged they are not easily marginalized. Conceptual (e.g. the commensurability of education systems), practical (e.g. sampling; translation), and technical (e.g. how data is aggregated and presented) issues are generally omitted. Including them would merely undermine the frame.

Step two: deal with issues

a) outsourcing and “established knowledge”

By attributing causal claims to other reports, or to “what we/everybody already know/s”, responsibility for dealing with issues is bypassed. This move often justifies the focus on an overarching policy lever in targeted studies, but may also be used to support specific claims. McKinsey’s (2010b) *Capturing the Leadership Premium* is illustrative. The *Leadership Premium* is a construct of the authors, implying the existence of a universally applicable, but elusive, formula for developing ‘effective’ educational leaders (determined by student outcomes).

The report begins by establishing the policy lever's centrality: 'officials in each of the systems [they] studied agree... [leadership] is crucial to outcomes' (5) and asserts the 'policy position is based on a *growing body of evidence* demonstrating the impact of effective school leadership' (5), which is 'consistent across a large number of countries and contexts' (5). Sources to support the claim are not provided, but precedents from other sectors further promote the policy lever's importance:

While... still a subject of debate in education, [leadership's] significance is now taken for granted in business, politics, the military, and almost every other area of public life. (6)

Across reports, claims such as 'the evidence suggests' (McKinsey 2010b, 12), or 'the evidence shows' (Schleicher 2009, 261), are often used but the source of the evidence is rarely provided. The strategy allows advocates to leverage their expertise, asserting their authority without revealing the source of their knowledge. In time, selected policy levers (such as 'leadership' or 'quality teaching') become axiomatic trends, falling into the realm of "what everybody knows". For example:

Hong Kong, Shanghai, Korea and Singapore all focus on the things that *are known to matter* in the classroom, including a relentless, practical focus on learning, and the creation of a strong culture of teacher education, research, collaboration, mentoring, feedback and sustained professional development. (Grattan Institute 2012, 2)

Though reasonable, no sources are provided to support these causal claims. The high-performing systems originally positioned as the focus of analysis are not used to develop the causal story, but are invoked to confirm what is *known to matter*. The selected practices are then bestowed with causal influence through their association with key levers that are *known* to have an influence.

When sources *are* provided, they are from the same genre, using common features, strong associations or what systems *tend to do* to identify 'best practices'. IPPR (2012) provides a model example:

As a result of international benchmarking, there is growing knowledge of how to reform education systems successfully... At the school level, this knowledge is set out in three major reports McKinsey's 2007 (Barber and Mourshed)... McKinsey's 2010 (Mourshed, Chijioke et al.)... and Marc Tucker's 2011 book. (58)

A table combining the knowledge contained in these reports is then presented with the title: 'The building blocks of world-class education systems: what we already know' (59). The McKinsey reports have been widely critiqued (e.g. Alexander 2010; Coffield 2012) and Tucker's 2011b book is an edited collection, providing descriptive overviews of high-performing systems. Greene (2012) points out that as the chapters follow no discernable method, we must 'trust the authority of the authors that they have correctly identified the relevant factors and have properly perceived the causal relationships'.

Tucker's report (2011a) is adapted from the final two chapters of his edited book (2011b), and is based on what *some*, *most* and *nearly all* systems do across a series of key levers:

What follows is a new agenda... derived from the experience of the countries that have consistently outperformed the United States. It was constructed simply by taking the subsection headings and reframing the language of the preceding sections in the form of an action agenda. (Tucker 2011, 40)

In short, non-specific sources are used to develop a descriptive (and partial) overview of 'what systems do' on selected policy levers. This portrayal of 'what *is*', is then reframed to advocate 'what *ought* to be done'. This is then referenced by IPPR (2012) to establish 'what we already know'. It is now apt to revisit the implications of this self-referencing, identified by Silova and Brehm (2015), whereby 'a hypothesis that has been tested in earlier research becomes cited as a proven and taken-for-granted reality... without critical engagement with alternative explanations' (27). In this way, speculative hypotheses are advanced and then inflated to the status of established knowledge in subsequent texts.

While broader influences on education outcomes may be dismissed outright when streamlining explanations (see *focus on factors 'amenable to control'*), they may also be dealt with when presenting recommendations. By *accepting and then rejecting limitations*, critique is pre-empted, portraying factors outside school as secondary concerns, unable to undermine the effectiveness of their 'best practices'. This strategy features prominently in each of the McKinsey reports (see Auld and Morris 2014, 144-145), for example:

It is important to recognize that there are contextual differences between systems, and that what works in one system may not work in another... the examples here should be taken as sources of insight and ideas, not as proven best practices which can be universally applied. *At the same time though, most of the evidence we have reviewed suggests that good leadership is the same irrespective of contexts and that "what works" is surprisingly consistent.* (McKinsey (2010b, 3 italics added)

The significance of context is then repeatedly undermined to support the causal story:

More than school inputs or context, learning depends on a determined and accountable school leader. (5)

This knowledge applies across a range of systems in different contexts. (7)

Although the leadership premium 'is a long way from being truly captured' (28), the conclusions are presented as 'what the consensus recognizes' (28). The 'consensus' refers to what 'some', 'most' or 'nearly all' system leaders and school leaders *believe* is important, presented as what they 'recognize'; i.e. the 'received wisdom' (what we *know*). Many of the conclusions, such as, 'systems policies and practices make a difference to leadership capacity' (28) are self-evident.

The Grattan Institute's (2012) dismissal of variables beyond its focus is followed with a presentation of policies pursued in *one or several* of the systems. The overview outlines six policy areas before asserting: 'no country can import another's culture, but these six programs have been the focus of reform in many systems throughout the world' (6). Here the issue of culture-context is acknowledged but dismissed, and causality is bypassed with a vague appeal to precedent in other systems.

Rather than reject or marginalise caveats directly, they may be circumvented, giving the impression that the advocate is respecting limitations while delivering the research ambition indirectly. This strategy involves three distinct sub-strategies (*hanging observations; softening the ambition; and, obfuscation*). Rather than make an explicit causal statement, attention is drawn to an apparent correlation, providing a *hanging observation* that leaves the reader to make the causal inference.

While the OECD (2012) explicitly rejects the possibility of identifying causal relationships, it outlines ways in which PISA can be used to improve education systems (22-23). One such application is as follows:

PISA can help governments to optimise existing policies or consider more fundamental alternatives when researchers combine *advanced forms of educational assessment with sophisticated survey research methods*... By linking these two bodies of data one can associate certain patterns of student performance with a multitude of background data... (23 italics in original text)

In this way, while the causal nature of such relationships might not be established, an extensive web of correlations can be drawn between certain dimensions of student performance and a large range of factors that *could conceivably affect* that performance. (23 italics added)

The report then appeals to industrial benchmarking techniques, and states:

The aim in this report is to relate differences in student achievement between one country and another to certain features of those countries' education systems. (23)

Elsewhere, the OECD (2013b) adopts a similar approach, highlighting a 'network of correlations' that can be used as 'a resource for decision making' (190). Valverde (2014) refers to this strategy as 'data inspired speculation' (486), a process which evokes Markkanen and Schroder's (1989) portrayal of scientific 'hedging', as any 'non-direct strategy of saying less than one means' (171). The practice is used to express uncertainty in propositions (Hyland 1998), thereby 'enable[ing] writers to refer to speculative possibilities while guarding against possible criticism' (443).

If the aim is merely to 'stimulate discussion' (OECD 2013b, 190), however, the actual policies highlighted (or their causal influences) are largely irrelevant, and the promotion of speculative explanations serves to focus education debate around the PISA data. This strategy of *loading* the discursive space is implicit in Schleicher's ready endorsement of studies that make questionable (and often conflicting) 'best practice' claims (e.g. McKinsey 2007; GEMS 2014). Rather than challenge the frame, researchers are drawn into speculative exchanges regarding the reasons for high

performance. Elsewhere, the OECD (2013a) contrasts the US with systems that perform well, or are rapidly improving:

While PISA cannot identify cause-and-effect relationships between policies/practices and student outcomes, it can show educators, policy makers and the interested public *how education systems are similar and different*. (14)

The strategy allows the response to any challenge: “we merely noted the common features, we neither attributed causality nor promoted policy transfer”. Indeed, when questioned on the OECD’s endorsement of the Common Core Standards in the U.S., Schleicher responded:

All we said was that Common Core Standards are in line with what they have in high-performing countries. We have empirical evidence for that. (Schleicher, quoted in The Wilby 2014, italics added)

Given the OECD’s increasing commitment to advocacy, such observations will inevitably be construed as prescriptions of ‘best practice’. Elsewhere, Schleicher (2011) has described the Common Core Standards as an ‘important step’ for raising outcomes in the United States. The strategy is continued in Schleicher (2014), which offers ‘pointers for policy and practice’. The report draws on PISA 2012 and other OECD reports (*outsourcing*) and asserts: ‘the analysis is complemented with examples that illustrate proven or promising practices in specific countries’ (17). How the identified practices have been *proven* to work is unclear, while a focus on *promising* practices bypasses the need for demonstrable impact.

Despite such ambiguity, the OECD’s veiled policy recommendations are widely cited as authoritative. The Grattan Institute (2012) illustrates how hanging recommendations are inflated into causal claims:

Two additional factors emphasise classroom management skills: the proportion of classroom time that is actually used for effective learning and teaching, and, school and classroom climate. The evidence shows that these are *universal qualities* of good teaching, and improve student learning. (16)

Two OECD reports (2008; 2010b) are cited as evidence. They highlight correlations, do not make universal claims, and do not make direct connections between the attitudes, beliefs or conditions identified by teachers in the surveys and actual improvement within the systems. Regardless, the claims are platitudes: more classroom time devoted to ‘effective’ learning would surely ‘improve’ student learning, and ‘good teaching’ is defined by its impact on student learning.

A variation is presented in *Worlds Apart* (1996), which states that it ‘represents an attempt not to test hypotheses... but rather to generate them’ (1). This *softening of the ambition* responds to the limitations and caveats, positioning the authors in the middle ground. As we have noted, ‘whereas testing hypotheses requires in-depth study and will likely produce complex explanations, the authors can generate hypotheses that are more easily translated into clear – if speculative – policy lessons’ (Auld and Morris 2014, 141). Though these hypotheses provide a basis for further investigation,

the *logic of action* (crisis, precedent, and derision) is employed to recommend ‘limited experimentation with non-British practice’ (59). These speculative hypotheses are still dredged up as ‘evidence’ to support reforms, including the National Curriculum in England (e.g. DfE 2011).

Rather than confront the issues directly, an implicit process of *obfuscation* is used to deliver the research ambition. Whether undertaking exploratory or targeted studies, explanations of how variables interact in a given context to produce the desired outcomes are avoided. Any attempt to identify how the effect of education practices or structures was established is concealed within an impenetrable black box. This is particularly apparent in the case of celebrity comparativists, who promote a coherent and consistent formula that is based on the lessons distilled from their extensive experience. As Tucker (2011) states:

After 22 years of research on the factors that account for the success of the countries with the best education record, I find myself convinced that seven things account for the lion’s share of the difference... (35)

Barber (2009) states that he is awake to the errors made by system leaders, ‘partly because I see them all around the world, but also because, at one time or another, I have made most of them myself’ (19). Again, we must have faith, and trust that the expert has identified the correct lessons from their experience. This issue is exacerbated in reports and projects involving multiple individuals, departments, or institutions, whereby tracing the processes and stages of analysis is hampered by complex layers of interaction and bureaucracy.

Move 5. Qualifying recommendations

Here we explore four strategies (*necessary conditions; knowledge is contextual; the burden of implementation; knowledge is imperfect*) that serve to qualify the status of policy recommendations, protecting the advocate from critique and deflecting responsibility if interventions are unsuccessful. These strategies are used to address the limitations and to provide a set of waivers. While the inclusion of these qualifications may be viewed positively, as an appropriate acknowledgement of the limitations, they are effectively used to co-opt critique, buttressing the policy core while using the *logic of action* to promote speculative causal claims as ‘best practices’.

Designating an intervention as a *necessary (but not sufficient) condition* to effect improvement is allied to the claim that it is a constant and transferrable principle. Some examples:

This focus on instruction, though a necessary condition, is in itself insufficient to bring about system improvement. (McKinsey 2007, 27)

In all these ways, the experience in the two cities (Singapore and Hong Kong) reflects the kind of reform in education that appears to be necessary and essential worldwide as the economy advances (OECD 2012, 171)

Describing interventions as *necessary (but not sufficient)* provides a powerful tool in the context of advocacy, not only implying that: (a) the selected intervention ‘works’ (and is transferrable); but also that, (b) the desired improvement *cannot be achieved* unless the identified *necessary* intervention is introduced. Flew (1975) clarifies the implications:

When such and such is being said to be not the logically but the causally necessary condition of this or that, then it is being said that, in the world as it actually is, with the laws of nature as they actually are, this or that could not in fact be produced without first producing or having such and such. (40-41)

Such deterministic claims are easily undermined by the production of a solitary counterexample, and these are readily available even in the array of explanations promoted by advocates. The strategy, a search for credibility ‘by appeal to notions culled from the philosophy of science’ (Barnes and Edge 1982, 246), is persuasive but always unfounded. Advocates may insist that interventions are *logically* necessary to improve educational outcomes. However, such analytical claims cannot be derived from, nor justified with reference to, international comparisons.

Second, the move allows advocates to promote policies while withholding guarantees of success. Failure can be attributed to the intervention not being enacted correctly, or failure to select the combination of interventions that together would be *sufficient* to effect the desired improvement (see *burden of implementation*). Finally, the strategy minimizes the likelihood of conflict with other advocates, who hold different secondary beliefs (i.e. policy preferences).

Though explanations are streamlined to omit culture and context, advocates stress that ‘best practices’ must be interpreted and enacted with sensitivity to the context (i.e. *knowledge is contextual*). This move is supported by the distinction between control and explanatory factors, and the posited existence of underlying ‘operational elements’. Flexible recommendations are thereby provided which marginalize caveats and protect the policy core by co-opting critique. For example:

Improving systems tailor how they implement the intervention cluster in each performance stage to their context. (McKinsey 2010a, 61)

By identifying the characteristics of high-performing education systems... PISA allows governments and educators to identify effective practices that they can then adapt to their local contexts. (OECD 2013a, 13; 2013b, 3)

How precisely the context was influential or how practices can be recontextualised is neither explored nor explained, and the onus for successful interpretation and implementation is located with local policymakers. Once ‘best practices’ have been abstracted from their original context, however, policymakers are left with an incomplete basis from which to begin any process of ‘translation’. More fundamentally, the move implicitly elevates speculative and provisional causal claims (e.g. hanging observations) to absolute status by shifting attention to the pragmatics of transfer.

Invoking *the burden of implementation* focuses responsibility for the success of interventions on governments and education leaders. For example, it may be asserted that an intervention failed because the selected policies were not *sufficient* to effect the desired change, or because they were not implemented or contextualized appropriately (e.g. McKinsey 2010a 26). Though the focus on enactment is clearly important, the strategy again elevates the status of claims, and effectively renders ‘best practices’ infallible. While gains in outcomes are used to demonstrate the effectiveness of interventions, blame for any failure to improve is located elsewhere within the system. Advocates can thereby continue promoting policies or education services as necessary in other contexts regardless of the outcomes in specific cases.

IPPR (2012) warns that ‘systems... need to become more adept at generating, identifying and scaling innovation internally’ (61), while the Grattan Institute (2012) identifies the elements for initiating behavioural change, *all* of which are required for successful implementation (21). It then notes the importance of ‘political will’, stating that ‘bureaucrats, teachers, parents and students may be fearful, and reluctant to change’ (22). If reforms fail, blame can first be directed towards a lack of political will, and then be shifted down onto one of the aforementioned ‘resistance groups’.

The OECD (2013a) repeatedly states the aim of investigating whether ‘*faithful* implementation’ of the Common Core State Standards for Mathematics (CCSSM) will raise PISA scores in the US, finally offering a qualified response: it is ‘intuitively plausible’ that ‘faithfully implementing the CCSSM would improve PISA results (90). If the projected outcome gains do not materialize, we can assume that the recommendation was not implemented faithfully, was confounded by resistance groups, or we were misled by our plausible intuitions. If successful, however, we might conclude that basing students’ learning on a specific test will raise their performance on that test.

Finally, the refrain that *knowledge is imperfect* both creates space to promote contrary recommendations in future reports and guards against negation. Advocates may note the imperfect nature of knowledge claims (e.g. McKinsey 2010a), present policies as ‘promising’ (OECD 2014), as ‘hypotheses’ meriting ‘experimentation’ (Reynolds and Farrell 1996), or make direct concessions. For example, ‘we still do not fully understand how to ensure consistent leadership across systems’ (McKinsey 2010b, 5). The McKinsey (2010a) report invokes pragmatism in its conclusion:, :

To begin with, our knowledge grows in spots. The spots may be large or small, but the knowledge never grows all over: some old knowledge always remains what it was. Your knowledge of pragmatism, let us suppose, is growing now. Later, its growth may involve considerable modification of opinions which you previously held to be true. But such modifications are apt to be gradual (122)

Any policymaker who finds that the ‘best practices’ used to inform interventions have been ‘updated’, can be reminded that knowledge is provisional and apt to change. As the McKinsey (2010a) report states:

Schools systems are constantly changing, so what worked a few years ago might well have little relevance today. (11)

The appeal provides the final element of *the logic of action*: (i) change is preferable to maintaining the status quo (see *crisis rhetoric*); and (ii) scientific knowledge is necessarily imperfect; therefore, (iii) it is most rational to pursue the required change based on current best conceptions of knowledge (i.e. ‘best practices’/‘what works’), even though causal explanations are imperfect and are apt (and likely) to change. The more compelling the crisis, the more likely a reader will accept knowledge claims that are not strictly warranted (Oxley, Vedlitz et al. 2014, 256). Appeals to precedent, the derision of non-believers, and an emphasis on actor expertise reinforce the logic.

Conclusion

We have identified a range of moves and strategies used to enhance the persuasiveness of texts advocating “best practices”, and specifically to overcome the issues of complexity and causality. The resulting framework, summarized in Table 2, is illustrative, and highlights the strategies that are used variously and in different combinations. Though advocates exhibit disparate backgrounds and agendas, and operate in diverse contexts, the analysis suggests that advocates follow a basic repertoire of common moves.

The *logic of action* establishes the necessity and viability of the task. By emphasizing their reputation or experience, or by demonstrating the scientific nature of comparisons and/or genre knowledge, authors establish their expertise and provide an authoritative basis for promoting claims. When handling limitations directly, a range of techniques streamline complexity, allowing the development of straightforward explanations that focus on factors amenable to control. Caveats included to demonstrate sensitivity to the issues are marginalized, and a variety of moves enable the delivery of the research ambition without making overt causal claims. Finally, speculative claims are recast as firm but flexible prescriptions, with contingencies placing the onus for successful interpretation and implementation elsewhere.

These strategies combine to create a system that is closed conceptually, collapsing complexity and reinterpreting the nature of social reality to enable the delivery of the research ambition. As Cowen (2014) observes, ‘the genie (i.e. context) is firmly back in the bottle’ (294). Though the strategies preserve the community’s deep core and policy core beliefs, the process of buttressing serves to hollow out comparisons, giving advocates a ready answer to any criticism while failing to address the underlying problems at its core. The result is the accumulation of decontextualized observations, elevated to the status of causal ‘facts’, the relative authority of which rests on perceptions of advocates’ expertise. Below we highlight the more generic issues that emerged from our analysis, which relate to: the process of discovery; the analytical nature of claims; and, the influence of the political context. Finally, we discuss the implications for the field.

First, just as the policy core is ensconced within an array of defensive fortifications, empirical anomalies are ignored or dismissed, allowing advocates to promote their preferred policies without fear of negation. This highlights the need to distinguish between how conclusions are developed and how they are justified. Two statements regarding the process of *discovery* are revealing:

The data is pretty raw stuff, isn't it? And it's open. There's normally more than one possible interpretation of data. So you bring your own perspective. (Barber 2014, 81)

Where [Schleicher] is brilliant, is to conclude. He is fantastic in this... He is the conclusions expert - they are in before the meeting... It is very convenient. (European Commission actor, quoted in Grek 2013)

Knowledge claims often follow established policy trends, or reflect actors' own experiences, professional-educational background and/or agenda. International examples are then cited as 'evidence', 'information selected from the available stock and introduced at a specific point in an argument' (Majone 1989, 48) to persuade the audience to accept the truth of a proposition. The practice reflects what Bloor (1982) refers to as 'the priority of the informal over the formal', whereby 'formal structures (of syllogisms) are connected to actual inferences by an interpretive process'. In such cases, 'formal logic represents a mode of display... a contrived and more or less artificial surface structure' used 'to strengthen and justify... predetermined conclusions by casting them in a deductive mode' (118-119). As Novoa and Yariv-Mashal (2003) note, 'the result is a "soft comparison" lacking any solid theoretical or methodological grounds' (425).

Second, many 'best practice' claims are analytical rather than empirical. When confronted with reports extolling the importance of *good* leadership, or a *high-quality* teaching profession - e.g. 'the performance of students is linked to the quality of teaching they receive' (Schleicher 2014, 11) - it is hard to argue for the antithesis, or null hypothesis: poor quality teachers improve learning outcomes. *High-quality* teaching and *good* leadership are defined by their relation to student outcomes, and so the statement tells us nothing more than can be derived from examination of the meaning of the words. As Hyslop-Marginson and Naseem (2007) argue with regard to other 'best practice' claims (e.g. the benefits of 'time on task', or 'student faculty contact'), 'none of these claims require empirical demonstration since they are simply matters of common sense and conceptual tautologies' (102).

Third, we acknowledge the context in which 'best practice' claims are developed. To be successful, advocates must develop authoritative claims that can be used to support reform directly. In short, "*without a straightforward policy story you are just another researcher with some data*". As universities continue to take on more of the characteristics and functions of commercial organisations, pressure to demonstrate impact and to secure funding has increased. Meanwhile, the status of international evidence in education means that the availability of funding is often conditional on the inclusion of a comparative element. Consequently, the community has an increasingly active membership in academic settings. These pressures are being

extended more concretely into the field of international development, signaled by the shift from “provision” to “quality” in the UN’s post-2015 goals.

PISA for Development will be anchored to the new targets, ranking systems by student outcomes on the selected competencies and allowing donor agencies to measure performance and track ‘progress’. Experts will emerge to identify ‘best (development) practice’ and thereby put each nation’s house in order, selling both their expertise and services to give systems a competitive advantage. System development will duly be oriented towards improvement on the chosen measurements, with any gains in outcomes used to demonstrate the legitimacy of the enterprise. As Rappleye (2010) asserts, ‘complexity will be increasingly collapsed according to the logic of power, political imperatives, *ideals* of ‘development’ and sources of funding’ (86). Once the goal of reform is reduced to improvement on PISA, the next logical step may simply be to align curricula and teaching with the assessments (Carnoy, Khavenson and Ivanova 2015).

Systematically relating national curricula and assessment to PISA is already stated as one of the survey’s possible applications (OECD 2012, 23). As countries increasingly begin to view their own education systems through the ‘prism of PISA’, we may anticipate further appeals to precedent, daring policymakers to ignore the trend. What was once denounced as ‘gaming’ will thus be recast as ‘benchmarking’, giving credence to the aphorism: *when you measure something, you change it*. A vision for the future of Comparative Education is laid out by the OECD (2012), a technical process modelled on industrial benchmarking, in which the outcomes have been determined, and the aim is simply to engage in the global war for talent by ‘learn(ing) enough from [our] competitors to beat them at their own game’ (24). The field must continue to analyse and challenge this ambition, strive to broaden rather than narrow research horizons, and remain open to the world beyond the prism of PISA.

Funding

This work was supported by The UK Economic and Social Research Council/Hong Kong Research Grants Council Joint Research Scheme under Grant ES/K010433/1.

References

- Alexander, R. (2010). "'World Class Schools' - Noble Aspiration or Globalised Hokum? BAICE Presidential Address at the 10th UKFIET Conference on Education and Development 2009." *Compare* 40(6): 801-817.
- Alexander, R. (2012). "Moral Panic, Miracle Cures and Educational Policy." *Scottish Educational Review* 44(1): 4-21.
- Auld, E. and P. Morris (2014). "Comparative Education, the 'New Paradigm' and Policy Borrowing: Constructing Knowledge for Education Reform." *Comparative Education* 50(2): 129-155.

- Ball, S. J. and S. Exley (2011). "Making Policy with 'Good Ideas': Policy Networks and the 'Intellectuals' of New Labour." *Journal of Education Policy* 25(2): 151-169.
- Barber, Kihn et al. (2011). *Deliverology: From Idea to Implementation*, London, McKinsey Education.
- Bhatia, V. (1993). *Analysing Genre: Language Use in Professional Settings*. London, Longman.
- Biesta, G. J. J. (2010). "Why 'What Works' Still Won't Work: From Evidence-Based Education to Value-Based Education." *Studies in Philosophy & Education* 29: 491-503.
- Barnes, B. and D. Edge (1982). Science as Expertise. In *Science in Context: Readings in the Sociology of Science*. B. Barnes and D. Edge. Milton Keynes, Open University Press: 233-249.
- Bloor, D. (1982). Formal and Informal Thought. In *Science in Context: Readings in the Sociology of Science*, edited by B. Barnes and D. Edge. Milton Keynes, Open University Press: 117-124.
- Carnoy, M., T. Khavenson, A. Ivanova (2015). Using TIMSS and PISA results to inform educational policy: a study of Russia and its neighbours. *Compare* 45(2): 248-271.
- Coffield, F. (2012). "Why the McKinsey Reports Will not Improve School Systems." *Journal of Education Policy* 27(1): 131-149.
- Coughlan, S. (2012). "China: the World's Cleverest Country?". Retrieved 05.06.2012, 2012, from <http://www.bbc.co.uk/news/business-17585201>.
- Cowen, R. (2014). Ways of Knowing, outcomes and 'comparative education': be careful what you pray for. *Comparative Education*, 50(3): 282-301.
- Dale, R. (2015). Conjunctions of Power and Comparative Education. *Compare*, 45(3): 341-362.
- Demszky, A. and A. Nassehi (2014). The Role of Knowledge in Scientific Policy Advice: Doing Knowledge. In *World Yearbook of Education 2014: Governing Knowledge*, edited by T. Fenwick, E. Mangez and J. Ozga. London, Routledge: 113-127.
- Fairclough, N. (2003). *Analysing Discourse: Textual Analysis for Social Research*. London, Routledge.
- Farnsworth, V. and Y. Solomon (2013). Introduction. In *Reframing Educational Research: Resisting the 'What Works Agenda*, edited by V. Farnsworth and Y. Solomon: 1-6.
- Feniger, Y. and A. Lefstein (2014). "How Not to Reason with PISA Data: an Ironic Investigation." *Journal of Education Policy*. doi: [10.1080/02680939.2014.892156](https://doi.org/10.1080/02680939.2014.892156)
- Fischer, F. (1993). Policy Discourse and the Politics of Washington Think Tanks. In *The Argumentative Turn in Policy Analysis and Planning*, edited by F. Fischer and J. Forester. USA, Duke University Press: 21-42.
- Flew, A. (1975). *Thinking about Thinking*. New York, Harpercollins.
- Goldstein, L. (1993). "The Fallacy of the Simple Question." *Analysis* 53(3): 178-181.
- Gove, M. (2011). *Michael Gove to the World Education Forum*. 11.01.2012. Retrieved from <https://www.gov.uk/government/speeches/michael-gove-to-the-education-world-forum>.
- Gove, M. (2013). "Oral Statement to Parliament: 2012: OECD PISA results." Retrieved from <http://www.gov.uk/government/speeches/2012-oecd-pisa-results>.
- Greene, J. P. (2012). *Best Practices are the Worst: Picking the Anecdotes You Want to Believe*. <http://educationnext.org/best-practices-are-the-worst/>

- Grek, S. (2013). "Expert moves: International Comparative Testing and the Rise of Expertocracy." *Journal of Education Policy* 28(5): 695-709.
- Hajer, M. A. (1993). Discourse Coalitions and the Institutionalization of Practice: The Case of Acid Rain in Britain. In *The Argumentative Turn in Policy Analysis and Planning*, edited by F. Fischer and J. Forester. USA, Duke University Press: 43-76.
- Hammersley, M. (2013). *The Myth of Research-Based Policy and Practice*. London, Sage.
- Henry, M., B. Lingard, F. Rizvi, and S. Taylor. 2001. The OECD, Globalization and Education Policy. Oxford: Pergamon Press.
- Henry, A. D. (2011). "Ideology, Power and the Structure of Policy Networks." *The Policy Studies Journal* 39(3): 361-383.
- Hopmann, S. T. and G. Brinek (2007). Introduction. In *PISA According to PISA: Does PISA Keep What It Promises?*, edited by S. T. Hopmann, G. Brinek and M. Retzl. Berlin, LIT Verlag.
- Hyland, K. (1998). *Hedging in Scientific Research Articles*. Amsterdam, John Benjamins Publishing Company.
- Hyslop-Margison, E. J. and A. Naseem (2007). *Scientism and Education: Empirical Research as Neo-Liberal Ideology*. New York, Springer.
- Jenkins-Smith, H., C. L. Silva, K. Gupta, and J. T. Ripberger (2014). "Belief System Continuity and Change in Policy Advocacy Coalitions: Using Cultural Theory to Specify Belief Systems, Coalitions, and Sources of Change." *The Policy Studies Journal* 42(4): 484-508.
- Jenkins-Smith, H., D. Nohrstedt, et al. (2014). The Advocacy Coalition Framework: Foundations, Evolution, and Ongoing Research. *Theories of the Policy Process*. P. A. Sabatier and C. M. Weible. Boulder, Westview Press: 183-224.
- Jones, M. D., M. K. McBeth, and E. A. Shanahan (2014). Introducing the Narrative Policy Framework. *The Science of Stories: Applications of the Narrative Policy Framework in Public Policy Analysis*. M. D. Jones, E. A. Shanahan and M. K. Mcbeth. Palgrave Macmillan, New York: 1-26.
- Kallo, J. (2009). *OECD Education Policy: A comparative and historical study focusing on the thematic reviews of tertiary education*. Jyvaskyla, Finnish Research Association.
- Kamens, D. H. and C. I. McNeely (2010). Globalization and the Growth of International Educational Testing and National Assessment, *Comparative Education Review*, 54(1): 5-25).
- Lachapelle, E., E. Montpetit, and J-P. Gauvin (2014). "Public Perceptions of Expert Credibility on Policy Issues: The Role of Expert Framing and Political Worldviews." *The Policy Studies Journal* 42(4): 674-697.
- Lemke, J., L. (1995). *Textual Politics: Discourse and Social Dynamics*. London, Taylor & Francis.
- Majone, G. (1989). *Evidence, Argument and Persuasion in the Policy Process*. London, Yale University Press.
- Markkanen, R. and H. Schroder (1989). Hedging as a Translation Problem of Scientific Texts. *Special Language: from humans thinking to thinking machines*. C. Lauren and M. Nordman. Clevedon, Multilingual Matters Ltd: 171-179.
- Mintrom, M. and P. Norman (2009). "Policy Entrepreneurship and Policy Change." *The Policy Studies Journal* 37(4): 649-667.

- ., P. (2012). "Pick n' Mix, Select and Project; Policy borrowing and the Quest for "World Class" Schooling: An Analysis of the 2010 Schools White Paper." *Journal of Education Policy* 27(1): 89-107.
- OECD (2008). *Creating Effective Learning and Teaching Environments: First Results from TALIS*. Paris, OECD Publishing.
- OECD (2010a). *The High Cost of Low Educational Performance: The Long-Run Impact of Improving PISA outcomes*. OECD Publishing.
- OECD (2010b). *TALIS 2008 Technical Report*. Paris, France, OECD Publishing.
- Oxley, D. R., A. Vedlitz, and B. D. Wood (2014). "The Effect of Persuasive Messages on Policy Problem Recognition." *The Policy Studies Journal* 42(2): 252-268.
- Ozga, J. (2014). "Knowledge, Inspection and the Work of Governing." *SISYPHUS: Journal of Education* 2(1): 16-38.
- Ozga, J. & M. Barber (2014). Data Work: Michael Barber in Conversation with Jenny Ozga. *Governing Knowledge*. T. Fenwick, E. Mangez and J. Ozga. London, Routledge: 75-86.
- Rapplee, J. (2010). Different Presumptions about Progress, Divergent Prescriptions for Peace: Connections Between Conflict, Development and Education in Nepal. In *Education, Conflict and Development*, edited by J. Paulson. Oxford, Symposium: 59-100.
- Rein, M. and D. Schon (1993). Reframing Policy Discourse. In *The Argumentative Turn in Policy Analysis and Planning*, edited by F. Fischer and J. Forester. USA, Duke University Press: 145-166.
- Robertson, S. L. (2011). *The Strange Non-Death of Neoliberal Privatisation in the World Bank's Education Strategy 2020*. Bristol, Centre for Globalisation, Education and Societies, University of Bristol.
- Sabatier, P. A. (1998). "The Advocacy Coalition Framework: Revisions and Relevance for Europe." *Journal of European Public Policy* 5(1): 98-130.
- Sabatier, P. A. and H. Jenkins-Smith (1993). *Policy Change and Learning: An Advocacy Coalition Approach*. CO, Westview Press.
- Schleicher, A. (2006). The Economics of knowledge: Why education is key for Europe's success. The Lisbon Council Policy Brief. Lisbon Council, Brussels.
- Silova, I. and W. Brehm (2015). From Myths to Models: The Reproduction of World Culture in Comparative Education. *Globalisation, Societies and Education* 13(1): 8-33
- Stewart, W. (2011). "The 'world's schoolmaster' delivers his annual report." Retrieved 24/05/2012, 2012, from <http://www.tes.co.uk/article.aspx?storycode=6112054>.
- Stone, D. A. (1988). *Policy Paradox and Political Reason*. USA, Harper Collins Publishers.
- Stone, D. A. (1989). "Causal Stories and the Formation of Policy Agendas." *Political Science Quarterly* 104(2): 281-300.
- Stone, D. A. (2012). *The Art of Political Decision Making*. New York, Norton & Company.
- Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge, Cambridge University Press.
- Tardy, C. M. (2011). Genre Analysis. *The Continuum Companion to Discourse Analysis*. K. Hyland and B. Paltridge. London, Continuum: 54-68.
- Valverde, G. A. (2014). "Educational Quality: Global Politics, Comparative Enquiry, and Opportunities to Learn." *Comparative Education Review* 58(4): 575-589.
- Waldow, F. (2012). Standardisation and Legitimacy: Two Central Concepts in Research on Borrowing and Lending. In *World Yearbook of Education 2012*:

Policy Borrowing and Lending in Education, edited by G. Steiner-Khamsi and F. Waldow. London, Routledge: 411-427.

Weinberg, A. M. (1974). "Science and Trans-Science." *Minerva* 10(2): 209-222.

Wilby, P. (2014) . Global School Tests under Attack as OECD Accused of Killing 'Joy of Learning'. *The Guardian*. online.

Yun, Y. and P. Morris (2015). Imagining School Autonomy in high-performing education systems: East Asia as a source of policy referncing in England, *Compare*, published online, DOI:10.1080/03057925.2015.1080115