Constructing School Autonomy with Accountability as a Global Policy Model: A Focus on OECD's Governance Mechanisms

Antoni Verger, Clara Fontdevila, and Lluís Parcerisa

INTRODUCTION

In the last decades, most countries in the world have faced major pressures to reform their educational systems. The emerging demand for global skills in increasingly interdependent economies, the challenges generated by technological innovation, and the competition between educational systems to achieve better results stand out among other sources of reform pressure. In this scenario, school autonomy with accountability (SAWA) reforms have been disseminated widely due to their promise to modernize education systems and strengthen their performance. In current educational reforms, school autonomy and accountability tend to be conceived as inseparable policy measures.

Governments (in their role as principals) are expected to give more autonomy to schools (the agents) in organizational, budgetary, and/or curricular terms—as long as schools accept being held subject to stricter supervision via external assessments and accountability measures. SAWA offers schools the possibility to adapt educational interventions to their local realities, thus giving more pedagogic and managerial powers to schools and strengthening the involvement of teachers (see also Holloway, this volume). By following the SAWA route, schools are expected to have the necessary room to maneuver to strengthen their instructional strategy and improve students' learning outcomes (De Grauwe 2005).

International organizations with great political reach, including the Organisation for Economic Co-operation and Development (OECD), have played a key role in the construction and diffusion of SAWA as a global policy model. During the 1980s, the OECD started promoting school effectiveness and school improvement research through the International School Improvement Project (ISIP) (Hopkins and Lagerweij 1996). The project was instrumental in galvanizing an expert network around school governance questions, and laid the foundations for the school improvement movement that took shape and consolidated during the

1990s (Bollen 1996; Hopkins and Reynolds 2001). The repertoire of policies advanced through these years prefigured to some extent the SAWA agenda that crystallized during the early 2000s, and particularly given their emphasis on schools as key units of change. Over the last two decades, the OECD has increasingly resorted to accountability, external assessments, and school autonomy measures when advising countries on how to organize and govern education (Bloem 2015; Niemann and Martens 2018). This shift has been paralleled by greater levels of conceptual and theoretical elaboration on the foundations and potential of these policy tools. For the OECD, SAWA is a policy model that seems to be context-resilient in the sense that it is expected to make education systems perform better in most territories, no matter their level of economic development or their administrative traditions. SAWA is a policy agenda that is congruent with the New Public Management (NPM) agenda that the OECD has advocated in different public sectors since the 1980s (Morgan 2009; Pal 2009). Both SAWA and NPM promote the fragmentation of public systems into smaller managerial units, outcomes-based management, and higher levels of accountability pressures among service providers.

This chapter aims at understanding the role of the OECD in the development and international dissemination of SAWA policies. Specifically, the chapter analyses the governance mechanisms through which these reforms are being promoted by the OECD, namely, data gathering, education policy evaluation, and the generation of policy ideas through different knowledge products and policy spaces. Methodologically, the chapter is based on a systematic literature review of a corpus of 33 papers, which we triangulate with official documents produced by the OECD. The chapter is structured as follows. In the first part, we present our research framework, which covers both our theoretical approach and our methods. In the second part, we present our main results, which we organize according to the different governance mechanisms articulated by the OECD around SAWA reforms. In the last part, we pick up the main points in a concluding discussion.

RESEARCH FRAMEWORK

The OECD as a Global Education Policy Actor

The OECD is an international organization (IO) that mainly operates as a platform for state cooperation and ideational exchange. The OECD describes itself as a "forum in which governments can work together to share experiences and seek solutions to common problems".[1] In comparison to other IOs, the OECD's legal and financial instruments are modest, and

its policy work mainly operates at the ideational level (Pal 2009). According to OECD officials, the main role of this IO in the education sector consists of comparing data, preparing policy reports for member countries, and facilitating horizontal learning between states (Schleicher and Zoido 2016). However, this is far from meaning the OECD is powerless, or that it simply operates as a neutral intermediary organization between states' interests and ideas. The OECD also plays an active role in agenda-setting, policy development, and policy transfer dynamics globally. In fact, the numerous knowledge-based activities developed by the OECD are a purposeful source of power, and a way for this IO to gain influence over its member (and even non-member) countries' policy agendas and decisions.

In the last decades, the OECD has become a key player in education, mainly via soft power mechanisms, and it is conceived as a prestigious interlocutor in the context of many education policy debates. The OECD's legitimacy in the education domain comes from its evidence-based and technocratic (i.e., apparently un-ideological) approach to problems of different natures, and its capacity to generate new sources of data and manage knowledge for policy purposes. According to Marcussen (2004, p. 29):

> the OECD is bound to play the so-called idea game through which it collects, manipulates and diffuses data, knowledge, visions and ideas to its member countries and, to a still larger extent, to a series of non-member countries.

As most IOs do, the OECD, since its formation in 1961, has focused on agenda-setting activities and, accordingly, has been telling countries the main problems they should prioritize in different policy sectors (Ougaard 2010). Nonetheless, more recently, this IO has become increasingly involved in policy development activities and its role has shifted from that   of problem identifier to that of solution builder. Over the years, the OECD

has built an extensive knowledge base, as well as the confidence and moral authority to develop and prescribe policy solutions (Berten and Leisering 2017). As we develop below, in education, the OECD's Programme for International Student Assessment (PISA) has become an inflection point in this respect, due to the capacity of this program to commensurate complex educational processes, such as teaching and learning, in concrete numerical indicators, and due to the country comparisons that derive from this quantification exercise (Martens 2007; Grek 2009; see also Lewis, this volume).

IOs can activate a range of governance mechanisms to transfer their preferred policy solutions to different territories and policy spaces. Martens and Jakobi (2010, p. 7) identify three main mechanisms of governance that are particularly present in the context of the OECD. They are: (1) Idea generation, defined as "a central activity by which the OECD stimulates political debates and develops new policy aims and goals"; (2) Policy evaluation, which "enables the organization to assess and guide a country's policy efforts"; and (3) Data production, which "although easily perceived as a non-political statistical exercise, makes it possible to compare countries directly with each other". As we show in the following sections, the OECD is very active and effective when it comes to triggering and articulating these three mechanisms with the goal of disseminating SAWA solutions internationally.

METHODS

Methodologically speaking, this chapter draws on the combination of two main data sources; namely, the results of a systematic literature review (SLR) on processes of education reform and policy instrumentation along the lines of SAWA; and a documentary analysis of OECD publications and working documents, with an explicit focus on accountability and school autonomy policies.

The SLR allowed for identifying clear instances of the OECD's influence over processes of reform and policy-shaping, as well as enabling circumstances and mechanisms that favor the penetration and appropriation of the OECD agenda in different contexts (see Verger et al. 2019). The review was informed by indexed publications retrieved through the SCOPUS database and followed the conventional steps of this literature review approach (cf. Gough et al. 2012). While the review contemplated a total of 158 papers, for our purposes here, we rely on a more reduced

subset of papers (*n* = 33) that inquired into the role of the OECD in SAWA reforms more explicitly.

The analysis of OECD publications, in turn, allowed us to characterize how educational problems and SAWA solutions are discursively articulated by this IO, and identify the causal ideas and principled beliefs through which the SAWA agenda is sustained and legitimized. We considered four main types of knowledge documents: documents published by the Centre for Educational Research and Innovation (CERI) with a focus on different aspects of education governance; country reviews of national policies for education; different documents concerned with PISA (including press releases and executive summaries of PISA reports); and working documents that included the agendas and/or main outcomes of meetings on SAWA-related issues held in the context of the OECD Directorate for Education and Skills.

OECD Governance Mechanisms and SAwA Reforms

Data Gathering

Data gathering is one of the main sources of IO power, since "the process of measuring, data production, and standard setting can make certain activities visible and legitimate and obscure other possibilities, conferring a self-propelling momentum on trends that may have been more consciously initiated" (Porter and Webb 2007, p. 48). The OECD has historically played an important role in "quantification and statistical mapping" in the field of education (Gorur 2015, p. 582). Quantification and data gathering allow the OECD to capture, describe, and compare the main characteristics of national education systems.

The OECD began to develop education indicators to promote the international comparison of educational systems in the 1960s. In the beginning, these indicators were intended to provide relevant information for educational planning to education ministers (Martens and Jakobi 2010). Nevertheless, in the mid-1960s, the education ministers of the member states pointed out to the OECD the need to produce comparable data in education in order to calculate the efficiency of education systems (Papadopoulos 1994). Since then, the OECD has developed variegated initiatives in this field, although it was not until the 1990s and the 2000s that this IO became the leading international organization (together with the World Bank) regarding the collection and production of comparable

statistical data in education (Cusso and d'Amico 2005; Sellar and Lingard 2013). PISA is the most influential data-intensive initiative of the OECD among political elites, in part because of its capacity of measuring and comparing learning achievement internationally, but is not the only one. Other well-known assessments and statistical products of this IO are Education at a Glance (EAG), the Programme for the International Assessment of Adult Competencies (PIAAC), or the Teaching and Learning International Survey (TALIS) (Gorur 2015; see also Lewis, this volume).

According to Martens and Niemann (2013), the mechanism of data gathering and production is crucial to understand the prominent role of the OECD in global governance. The OECD cannot sanction member countries per se, but the "naming and shaming" dynamics it promotes through the generation of league tables, such as those included in the PISA reports, are an effective form of framing and conditioning country decisions and promote the engagement of countries in a sort of "global education race" (cf. Sellar et al. 2017). This educational race intensifies for political but also economic reasons since, in a highly competitive and interdependent economic environment, learning achievement becomes a governmental asset to attract foreign investors and to aspire to generate more knowledge-intensive jobs.

PISA has had a substantive effect in the promotion of SAWA policies at the national level. The release of PISA results are central to many governments' perception about the education quality, equity, and/or efficiency problems that they face, and triggers educational debates of different intensities in countries and regions. As a result of these debates, and the consolidation of a narrative of a "learning crisis" that accompany these debates, policy-makers are inclined to select policy solutions that allow them to keep a better control of their educational results, such as national assessments and test-based accountabilities. In countries such as Germany, Switzerland, England, Denmark, Australia, Spain, and Norway, PISA results have fostered public debates leading to the adoption of SAWA policies at some level (Baxter and Clarke 2013; Engel 2015; Gorur 2015; Møller and Skedsmo 2013; Moos 2014; Sjøberg 2016; Sellar and Lingard 2013).

For example, in Germany, the so-called PISA shock in the early 2000s triggered the adoption of an output-oriented governance approach (Niemann et al. 2017), while in Norway, the scandalization triggered by the publication of PISA 2000 and 2003 was key to build a political consensus around accountability and quality assurance systems (Hatch 2013;

Camphuijsen et al. 2018). In other countries, such as Spain, PISA has been used strategically to legitimize the adoption of large-scale assessments and managerial school autonomy (Engel 2015; Popp 2014; Verger and Curran 2014), and has enabled bipartisan convergence around NPM policy principles (Dobbins and Christ 2019). In the case of England, mediocre PISA results in the first editions did not result into a shock, but were used to problematize the role of the inspection agency and, specifically, to promote a more instructional improvement approach among inspection services (Baxter 2014).

The statistical data produced by the OECD, mainly via PISA, triggers competition dynamics at both international and national levels. In the context of the so-called education race generated by PISA, the enactment of national large-scale assessments and accountability policies has become very strategic, and a necessary condition for governments to promote better learning outcomes and activate school improvement dynamics at a distance. By promoting competition, PISA reinforces the adoption of SAWA policies indirectly. However, the OECD, via PISA, also advocates these types of policies explicitly. In numerous PISA reports, SAWA is portrayed as an appropriate way to address performance issues and improve learning outcomes. As stated in the press release of PISA 2006:

> more importantly, there are a number of school policies and practices that are crucial for performance without being necessarily tied to resources. Let me just highlight three of them—institutional differentiation, autonomy, and accountability, because they feature so prominently in national education policy debates. (OECD 2007, p. 1)

The performative effects of international rankings can be observed in Denmark, where, according to Moos (2014), the prime minister stated in 2010 that the aim of the educational system was to become one of the top-five nations listed in the PISA report. In developing countries and emerging economies, the country aspirations might be more modest, but they also use PISA as a benchmark in their educational development plans. For example, in the case of a non-OECD country such as Brazil:

> The explicit goal of the PDE[2] in 2007 was for the IDEB[3] score to reach the OECD average level by 2021 (INEP, 2016; PDE, 2007), thereby achieving full alignment with international large-scale assessment models. (Kauko et al. 2018, p. 570; see also Toledo Silva, this volume)

PISA is influential not only due to the data it gathers and how it presents it, but also because it has become an "international reference" in terms of assessment systems. As an international reference, PISA has contributed to the reconfiguration of national assessments at the country level. Countries as diverse as Brazil (Kauko et al. 2018), Spain (Verger et al. 2018), Chile (Parcerisa and Falabella 2017), and Canada (Martino and Rezai-Rashti 2013) have adopted national assessments in the image of PISA, or have adjusted their national assessments to PISA methodology. The strategic adaptation of existing national assessments to PISA methodology, contents, and standards is seen as a way to modernize the national assessment system, but also of strengthening countries' competitiveness in the global educational race.

Overall, PISA has reached a significant impact and popularity among OECD member and non-member countries, and the data produced by this IO through PISA enjoys great credibility and legitimacy globally (Gorur 2015). As we show in the next sections, the ideational environment created by the comparative and audit culture (cf. Kamens 2013) that OECD induces via PISA facilitates also the deployment of idea generation, meaning-making, and persuasion strategies, which on many occasions result in the diffusion of SAWA policies internationally.

Policy Evaluation

The second OECD governance mechanism considered by Martens and Jakobi (2010) is policy evaluation, that is, activities oriented at providing countries with education guidance and policy advice. In the context of IOs, policy evaluations usually adopt the form of peer reviews in which experts in a particular policy domain identify policy issues and put forward policy recommendations. Peer review has been characterized as a practice revolving on multilateral surveillance and indirect coercion (Marcussen 2004; Porter and Webb 2007) that sets in motion benchmarking dynamics (Schuller 2005). In fact, the OECD is explicit on the fact that the effectiveness of its reviews relies on the peer pressure resulting from informal dialogues, public scrutiny, comparisons, and the impact of public opinion.[4] As noted by Pagani (2002), while peer review is not exclusive to the OECD, this IO has made a distinctly extensive use of this monitoring technique, to the point that it features today as one of its most distinctive practices (see also Krejsler, this volume).

In the OECD's education work, the most extensively used and prominent evaluation instruments are the Reviews of National Policies for Education (RNPEs). Other related instruments are Thematic Reviews, Country Background Reports, Country Notes, Country Case Studies, and Education Policy Outlooks. Policy development initiatives of the OECD, such as *Synergies for Better Learning*, also rely on background country peer evaluations. While such evaluation modalities differ in procedures and form, they all are oriented at identifying the country's weaknesses and main educational challenges, and suggesting and prioritizing certain courses of action under the forms of policy recommendations. RNPEs, however, stand out on the grounds of their interactive nature, their reliance on faceto-face interactions and field visits, and the length of the policy cycle they trigger off (e.g., the frequency and nature of follow-up mechanisms). For all these reasons, RNPEs are instrumental in the diffusion of the OECD educational agenda, the internalization of such agenda among national policy-makers, and the consolidation of the OECD as a policy expert and authorized knowledge producer (Grek 2017).

OECD country reviews have been strategic in the advancement of SAWA reforms in several countries, although not always for the same reasons. In some cases, country reviews are one of the sources that have contributed to problematize some aspects (or the general functioning) of the education system, and trigger domestic dynamics similar to those sparked by PISA results (see above). This has been the case of Scotland (UK), where the OECD report published in 2007 (*Quality and Equity of Schooling in Scotland*) animated a policy debate on the organization and management of the schooling system. Although the review recommendations were hardly observed by the new Scottish National Party government that took over in 2007, Grek (2016) draws attention to the crucial effect that the report had in changing both the parameters of the national debate and the country's self-perception.

In other cases, country reviews have a clearer or more direct effect in the selection of new policy instruments or the redefinition of existing ones. This is, for instance, the case of Chile, where, amid a national debate about the quality of education, the OECD country review published in 2004 was instrumental in securing the accountability solution. The OECD report not only affected the general direction of policy-making efforts, but also translated into a series of changes in the design and procedures of the national standardized test (SIMCE) (Bravo 2011; Parcerisa and Falabella 2017).

The case of Norway is especially illustrative of how the OECD's evaluation practices influence both the identification and prioritization of particular policy problems, and the selection of particular policy responses. A series of reviews and reports published by the OECD during the last 20 years in Norway have shaped the terms of the national education debate and have contributed decisively to the advancement of a system of low-stakes accountability. As early as in 1988, a RNPE conducted by the OECD questioned the high levels of decentralization of the education system, and called for the need to introduce centralized control mechanisms. Later on, in the mid-2000s, another OECD review on *Lifelong Learning in Norway* also played a significant role in securing the centrality of assessment and outcomes within the new educational strategy devised in the Green Paper (2003) and White Paper (2004), accordingly (Baek et al. 2018).

However, the publication of a country review or the engagement of a given country in the review process cannot be mechanically associated to policy changes or the reconfiguration of education agendas. As noted by Schuller (2005, p. 177), "the utility of the national reviews depends in part on the willingness of the country to confront issues and to be candid in the information it supplies"—and such readiness cannot be taken for granted. While OECD reviews in some contexts have provided a turning point, their role is much more limited in other settings or circumstances. Likewise, the impact of a review does not always last, nor does it entail a durable change in the framing of the debate and national priorities. In Ireland, for instance, the shift toward a more managerial or economistic approach to education was evident in the 1992 Green Papers that were themselves largely inspired by the OECD's RNPE of 1991 (Halton 2003). However, this approach was soon eventually abandoned following the appointment of a new Minister of Education and the negative consequences of this reform approach (with an excessive focus on accountability and rankings) observed in England (Halton 2003).

To be sure, the effect of evaluation practices in terms of policy change does not depend exclusively on their impact over decision-makers. The ultimate impact of OECD reviews depends also on their public reception, and on how the public puts pressure on governments according to the review findings (Porter and Webb 2007). The Chilean case is particularly illustrative of such dynamics. As documented by Parcerisa and Falabella (2017), the OECD review published in 2004 helped frame the discourse of the student movement, which was highly critical of segregation and the

inequality induced by the quasi-market system. These protests called into question the education system and, along with poor results in student performance, contributed decisively to create a context favorable to the advancement of accountability reforms. However, such dynamics do not appear to be the rule, but the exception. In fact, reviews are likely to receive only limited attention within domestic constituencies. On occasions, the government might even act as a gatekeeper and limit dissemination practices or media attention, thus neutralizing a key influence channel. So, neither the impact of OECD reviews over national policies, nor their effect over public opinion and public sentiments, can be taken for granted. In most cases, the transformative potential of OECD reviews, and their capacity to bring about a process of policy change, remains an empirical question.

However, and regardless of the direct effects of OECD reviews in terms of national policy change, another of the effects of this evaluation labor is the consolidation and dissemination of a common understanding of what constitutes appropriate policy. As noted by Porter and Webb (2007), "peer review requires standards and criteria against which a member state's policies can be reviewed" (p. 6), and such criteria are not always explicit. Rather, policy evaluation is likely to be guided by implicit principles and standards. This is why, even if their ultimate impact on domestic policy-making is uncertain, country reviews and RNPEs are informative in their own right.

A look at the RNPEs published in the last decade suggests that the recommendations about accountability included in these reviews, despite being tailored to country specificities, tend to respond to a few common principled beliefs. Among them, two principles feature prominently. First, one can see the need to encourage or consolidate a shift from inputs and procedures to outcomes as a focus of policy attention. This is particularly so in middle-income countries, such as Brazil (see OECD 2010) or Costa Rica (see OECD 2017), where an input-centered perspective has long prevailed. Second, the need to instill an evaluation culture predicated on greater levels of accountability and transparency, and involving the appraisal of different education stakeholders (including teacher and schools but also local and national administrators) against well-defined performance and quality benchmarks and standards. Such ideas can indeed be found in most of the reviews conducted in the last decade,[5] and are explicitly positioned as drivers of high-performing education systems (see, for instance, the reviews recently conducted in Norway [OECD 2011a] or

Greece [OECD 2018]). Overall, accountability and autonomy are generally emphasized as part of the necessary strategy to address broadly defined goals and portrayed as intrinsically desirable, but less clearly framed as a response to a particular challenge.

When it comes to specific policy instruments and tools, RNPE appear to be more heterogeneous in nature. Also, the degree of precision and detail in the definition of instrument settings is also quite variable. For instance, while calls for the strengthening of monitoring mechanisms and teacher and school appraisal schemes are widespread, the stakes attached to such assessments are subject to variation and are not always clearly delimited. A number of the reviews are, in fact, rather ambiguous or vague on the desirability of these arrangements—and some even call for some caution in their implementation. This is the case of the RNPE conducted in Colombia, which calls into question the design of the (collective) teacher bonus currently in place and warns against attaching increasingly higher stakes to a single measure, at the same time that it admittedly remains vague on what could constitute a more desirable accountability scheme (see OECD 2016). In general, in RNPEs, it is possible to identify a general preference for complex quality assurance systems relying on a wide range of sources, even if recommendations regarding specific policy- design features are more heterogeneous and subject to great variation in terms of the detail provided.

Idea Generation

Idea generation is a broad governance mechanism by which the OECD promotes policy debates between countries and constructs policy recommendations and programs. Two main types of policy ideas are being generated and disseminated in the context of IOs such as the OECD: causal ideas and principled beliefs. Causal ideas are more scientific and evidence- based in nature (i.e., *what works*), whereas principled beliefs respond to a more normative approach to policy (i.e., *what is acceptable*) (Marcussen 2004). Both types of ideas are connected in the sense that the tested policies, best practices, and benchmarks that the OECD promotes need to fit within its "explicitly normative agenda" (Pal 2009, p. 1061).

Idea generation is connected to—and, to a great extent, grounded on—the two governance mechanisms mentioned above (i.e., statistical data generation and policy evaluation), but it involves the proactive interpretation of these data and evaluations for policy purposes. The OECD is

well aware of the policy potential of the data it generates, especially through successful initiatives such as PISA. In fact, initially, the OECD outsourced the elaboration of the PISA report to independent consultants. However, it stopped doing so with the 2006 PISA cycle since, for the OECD, producing the PISA reports in-house is the best way of framing and controlling the policy message that derives from this international large-scale assessment (Bloem 2015).

PISA is the most important source for the policy development activities of the OECD. *PISA in Focus, Education Indicators in Focus*, and the *Strong Performers and Successful Reformers* video series are good examples of the OECD attempt of interpreting and translating PISA's quantitative results into tangible policy lessons and recommendations (Bloem 2015). SAWA recommendations are consistently present in these knowledge products, but also in other Directorate for Education and Skills initiatives. These include the *Synergies for Better Learning* series (Bloem 2015), or CERI initiatives such as *Governing Complex Education Systems* and *Strategic Education Governance*, which involve the participation of country representatives in a series of seminars and meetings, during which good practices are identified and the potential and weaknesses of country participants are assessed.[6] Beyond written products and international seminars, another way the OECD disseminates policy ideas is by giving advice to governments. Several studies show that the OECD has played a direct role as advisor in the context of SAWA reforms in countries like Spain (Dobbins and Christ 2019; Engel 2015; Verger and Curran 2014), Portugal (Carvalho and Costa 2017), and Mexico (Echávarri and Peraza 2017). In the context of these interactions at the country level, OECD officials play a strategic role in advancing educational reform by establishing bridges between international discourses and national politics, and by mediating between the fields of policy and research (Carvalho and Costa 2017).

The SAWA agenda gained centrality in the OECD internal discussions in the middle of the 2000s. In 2004, the OECD education ministers' meeting debated test-based accountability issues and "revealed wide differences of view among countries in how information on student learning outcomes can and should be used" (OECD 2005, p. 3). This debate triggered a series of informal meetings, training, and research initiatives with external experts and country representatives. The main focus of these initiatives was to find out whether increased accountability could benefit students' achievement and make educational results more equitable. The research base of these lines of work was to a great extent developed by the

German economist Ludger Woessmann. This author, after conducting different analysis with the PISA database, concluded that "enhanced accountability is associated with overall improvements in performance", and that there are "important complementarities between some testing and aspects of school autonomy as well as parental choice" (OECD 2008, p. 7).

The main policy message that derives from this line of work on SAWA, and which has been echoed in different OECD publications like *PISA in Focus. No. 9* (see OECD 2011b), is that managerial school autonomy is conducive to better educational results when combined with school rankings. Despite these findings being not necessarily conclusive (or could be subject to multiple interpretations), they have provided with arguments to advocates of SAWA when promoting educational reforms at the country level (see for instance Pagès and Prieto 2019, forthcoming). Nonetheless, to be fair, initiatives of the OECD on evaluation and assessment that came in the 2010s are more nuanced in their conclusions. These more recent initiatives do not advocate rankings and market accountability so strongly, but rather encourage formative assessments and multi-stakeholder accountability systems that do not put pressure exclusively on school actors (see, for instance, Burns and Köster 2016; OECD 2013).

Through the promotion of SAWA, the OECD advocates something else than a specific policy program or specific policy instruments. Rather, it promotes broader policy principles and a normative understanding of how educational systems should be modernized and organized. The OECD policy work has a strong normative dimension and, as it has been acknowledged by Woodward (2009, cited in Sellar and Lingard 2013, p. 715), "it is arguably through challenging and changing the mindsets of the people involved that the [OECD] achieves its greatest influence".

In education, the normative work of the OECD has contributed to governments taking for granted that they need to adopt assessment instruments, accountability measures, and performance targets to develop modern educational systems. Through initiatives such as PISA, the OECD has also contributed to governments assuming that datafication and a comparative perspective to educational performance are both effective and appropriate governance mechanisms that can improve educational systems in different ways. Several studies on SAWA reforms point to the role of the OECD in the promotion of an "evaluation culture" that, in countries such as Italy (Grimaldi and Serpieri 2014) and Ireland (McNamara et al. 2009), has contributed to the disruption of the conventional organization of edu-

cation systems, and has altered the ideational context of educational reform. This literature implies that some governments would be adopting national assessments and accountability systems by following a logic of appropriateness, and not necessarily a logic of consequences. According to McNamara and colleagues (2009), the adoption of the Whole School Evaluation program in Ireland, and the emerging evaluation culture in the country, did not seem to respond clearly to any particular problem. It was, to some extent, a response to demands from employers' groups and sections of the media favorable to greater levels of accountability, but it was mainly motivated by the need of the Irish government to "comply with" EU and OECD policy recommendations.

CONCLUDING DISCUSSION

The OECD's role in policy transfer dynamics goes beyond that of a neutral broker between states. This IO develops and actively disseminates policy ideas, such as school autonomy with accountability (SAWA), in the education sector, which fit within the broader agenda of the OECD on public sector administration reform. To a great extent, the SAWA agenda is congruent with the premise that, for the OECD, "it is vitally important to constantly try to improve [public sector] management practices and institutions" as a way to promote "the economic success and competitiveness" of its member countries (Pal 2009, p. 1078). SAWA provides the OECD with a narrative about educational reform that focuses on school governance instruments, but that is wide enough to accommodate different political perspectives on how to use these instruments and for which purposes.

To disseminate SAWA, the OECD activates three main mechanisms of soft governance; namely, data gathering, policy evaluation, and idea generation. Figure 11.2 summarizes the main initiatives through which the OECD mechanisms operate, and refers to a series of countries where these mechanisms have been documented as influential. While the figure refers only to OECD member countries, the reliance on soft governance mechanisms (and the fact that participation in policy evaluation and data gathering practices is not restricted to member states) enables the OECD to extend its influence beyond its membership (see Lewis, this volume).

	Data gathering	Policy evaluation	Idea generation
Mechanism of influence	Competitive dynamics at the international and national levels Identification of policy problems/agenda-setting Creation of political opportunity windows to advance/legitimize policy change	*Evaluation process*: Internalization of causal and principled beliefs among national decision-makers *Evaluation results*: Identification of policy problems/agenda-setting Identification and shaping of policy solutions	Dissemination of causal and principled beliefs through persuasion and technical advice Advancement of a comparative perspective encouragement of normative emulation practices
Main products or initiatives	PISA Education at a Glance TALIS PIAAC	Reviews of National Policies for Education, Thematic Reviews, Country Background Reports, Country Notes and Country Case Studies, and Education Policy Outlook	PISA in Focus, the Strong Performers and Successful Reformers series, Strategic Education Governance
Illustrative country cases	USA, Germany, England, Spain, Norway, Denmark	Scotland, Norway, Chile	Spain, Portugal, Mexico, Italy, Ireland

Fig. 11.2 OECD governance mechanisms as used in the promotion of SAWA reforms

It should be noted, however, that even if certain governance mechanisms play a particularly prominent or visible role in some countries, they never operate in a void. Rather, different governance mechanisms reinforce each other. It is precisely when these mechanisms are activated simultaneously that the SAWA proposals of the OECD are more likely to prove influential at the country level. For instance, the fact that a policy review of the OECD comes after a PISA scandal is expected to make countries more receptive to this IO's reform recommendations.

The OECD governance mechanisms over domestic policy do not necessarily operate vertically. OECD reviews lend themselves to countries' appropriation and instrumentalization in a particularly clear way, as a consequence of its *à la carte* and interactive nature. Hence, governments that commission a national review are likely to use the review to legitimize or advance their own agenda (Schuller 2005). Similar bottom-up dynamics can be observed in the case of the publication of PISA results, which governments interpret in such a way that better serves their political interests and/or policy preferences. Such instrumentalization dynamics should not be understood as deviant behavior on the part of evaluated states. Rather, as advanced by Centeno (2017), the OECD is frequently "eager to be 'instrumentalized'" (p. 100), as the OECD agenda is only likely to succeed, acquire legitimacy, and enjoy wide circulation if its proposals resonate with national interests (see also Toledo Silva, this volume).

Overall, the potential for the OECD governance mechanisms to advance the SAWA agenda appears to lie in their capacity to open a policy window through which the problem, policy, and politics streams (cf. Kingdon 1984) are affected in a relatively coordinated and coherent way. Through national reviews, data gathering, and idea generation initiatives, the OECD effectively acts as an *instrument constituency* (cf. Béland and Howlett 2016) that is able to theorize and operationalize SAWA policy instruments, at the same time as it matches these instruments to a wide range of problems. As an instrument constituency, the OECD promotes SAWA solutions in very different settings and attaches these solutions to a broad range of problems, including lack of transparency in public administration, low overall performance of the educational system, equity issues and learning gaps, lack of teachers' engagement, and so on. However, the OECD is more ambivalent on the particular uses and configurations of the different SAWA instruments and, in the case of national policy reviews and other types of country documents, it tends to adjust its more concrete SAWA prescriptions to contextual specificities.

Future research should analyze how the SAWA agenda has evolved over time within the OECD, and more systematically look at the role of this IO in the promotion of SAWA reforms. Assessing the OECD's influence on national policy-making is methodologically challenging, due to the fact that the kind of changes put forward by this IO are not readily observable. Evaluation practices, regular encounters, and other forms of exchange trigger dynamics of normative suasion powered primarily by socialization practices, face-to-face interactions, and meaning-making exchanges are particularly difficult to trace and document empirically (Checkel 2005; Grek 2017). As noted by Beyeler (2004), congruence between the OECD policy ideas and policy change is not enough to establish proof of the OECD influence, and references to the OECD in the context in national debates are an equally unsatisfactory proxy. The OECD footprints over processes of national policy reform remain thus an elusive, as well as fascinating, subject for future research.

NOTES

1. See www.oecd/about.
2. Plano de Desenvolvimento da Educação (or Education Development Plan).
3. Index of Basic Education Development.
4. Cf. https://www.oecd.org/site/peerreview/peerpressurearelatedconcept.htm. See also Pagani (2002).
5. It should be noted, however, that this agenda appears to be less clearly defined or well articulated for reviews published during the late 1990s and early 2000s. In these RNPEs, calls to accountability, external evaluation, and monitoring are sparser, and more unequally distributed, among reviews and regions.
6. See, for instance, http://www.oecd.org/education/ceri/strategic-education-governance.htm.

REFERENCES

Arnott, M. (2014). Westminster no more? Governance, devolution and delegation in Scotland. *Local Government Studies, 40*(6), 910–922.

Baek, C., Hörmann, B., Karseth, B., Pizmony-Levy, O., Sivesind, K., & Steiner-Khamsi, G. (2018). Policy learning in Norwegian school reform: A social network analysis of the 2020 incremental reform. *Nordic Journal of Studies in Educational Policy, 4*(1), 24–37.

Baxter, J. A. (2014). An independent inspectorate? Addressing the paradoxes of educational inspection in 2013. *School Leadership & Management, 34*(1), 21–38.

Baxter, J., & Clarke, J. (2013). Farewell to the tick box inspector? Ofsted and the changing regime of school inspection in England. *Oxford Review of Education, 39*(5), 702–718.

Bollen, R. (1996). School effectiveness and school improvement: The intellectual and policy context. In D. Reynolds, R. Bollen, B. Creemers, D. Hopkins, L. Stoll, & N. Lagerweij (Eds.), *Making good schools. Linking school effectiveness and school improvement* (pp. 1–19). London: Routledge.

Béland, D., & Howlett, M. (2016). How solutions chase problems: Instrument constituencies in the policy process. *Governance, 29*(3), 393–409.

Berten, J., & Leisering, L. (2017). Social policy by numbers. How international organisations construct global policy proposals. *International Journal of Social Welfare, 26*(2), 151–167.

Beyeler, M. (2004). Introduction: A comparative study of the OECD and European welfare states. In K. Armingeon & M. Beyeler (Eds.), *The OECD and European welfare states* (pp. 1–12). Cheltenham & Northampton: Edward Elgar Publishing.

Bloem, S. (2015). The OECD directorate for education as an independent knowledge producer through PISA. In H. G. Kotthoff & E. Klerides (Eds.), *Governing educational spaces* (pp. 169–185). Rotterdam: Sense Publishers.

Bonal, X., & Tarabini, A. (2013). The role of PISA in shaping hegemonic educational discourses, policies and practices: The case of Spain. *Research in Comparative and International Education, 8*(3), 335–341.

Bravo, J. (2011). SIMCE: Pasado, presente y futuro del sistema nacional de evaluación [SIMCE: Past, present and future of the national system of evaluation]. *Estudios Públicos, 123*, 189–211.

Breakspear, S. (2012). *The policy impact of PISA: An exploration of the normative effects of international benchmarking in school system performance*. OECD Education Working Papers, No. 71. Paris: OECD Publishing.

Burns, T., & Köster, F. (2016). *Educational research and innovation governing education in a complex world*. Paris: OECD Publishing.

Camphuijsen, M., Skedsmo, G., & Møller, J. (2018). School autonomy with accountability as a global education reform: Its adoption and re-contextualization in the Norwegian context. Paper presented at the European Conference on Educational Research (ECER), September 2018, Bolzano, Italy.

Carnoy, M., & Rhoten, D. (2002). What does globalization mean for educational change? A comparative approach. *Comparative Education Review, 46*(1), 1–9.

Carvalho, L. M. (2012). The fabrications and travels of a knowledge-policy instrument. *European Educational Research Journal, 11*(2), 172–188.

Carvalho, L. M., & Costa, E. (2017). Avaliação externa das escolas em Portugal: Atores, conhecimentos, modos de regulação [External evaluation of schools in Portugal: Actors, knowledges and regulation modes]. *Revista Brasileira de Política e Administração Da Educação, 33*(3), 685–705.

Centeno, V. G. (2017). *The OECD's educational agendas – Framed from above, fed from below, determined in interaction. A study on the recurrent education agenda*. Berlin: Peter Lang.

Checkel, J. T. (2005). International institutions and socialization in Europe: Introduction and framework. *International Organization, 59*(4), 801–826.

Cusso, R., & d'Amico, S. (2005). From development comparatism to globalization comparativism: Towards more normative international education statistics. *Comparative Education, 41*(2), 199–216.

De Grauwe, A. (2005). Improving the quality of education through school-based management: Learning from international experiences. *International Review of Education, 51*(4), 269–287.

Dobbins, M., & Christ, C. (2019). Do they matter in education politics? The influence of political parties and teacher unions on school governance reforms in Spain. *Journal of Education Policy, 34*(1), 61–82.

Echávarri, J., & Peraza, C. (2017). Modernizing schools in Mexico: The rise of teacher assessment and school-based management policies. *Education Policy Analysis Archives, 25*(90), 1–30.

Engel, L. C. (2015). Steering the national: Exploring the education policy uses of PISA in Spain. *European Education, 47*(2), 100–116.

Gough, D., Thomas, J., & Oliver, S. (2012). Clarifying differences between review designs and methods. *Systematic Reviews, 1*(28), 1–9.

Gorur, R. (2015). Producing calculable worlds: Education at a glance. *Discourse: Studies in the Cultural Politics of Education, 36*(4), 578–595.

Grek, S. (2009). Governing by numbers: The PISA effect in Europe. *Journal of Education Policy, 24*(1), 23–37.

Grek, S. (2016, September). *Beyond the standardisation vs. contextualisation debate: The role of the OECD in European education governance*. Paper presented at the International Organizations and the Globalization of Public Instruments and Ideas (IOGLOB) Seminar, Strasbourg.

Grek, S. (2017). Socialisation, learning and the OECD's reviews of National Policies for education: The case of Sweden. *Critical Studies in Education, 58*(3), 295–310.

Grimaldi, E., & Serpieri, R. (2014). Italian education beyond hierarchy: Governance, evaluation and headship. *Educational Management Administration and Leadership, 42*(4S), 119–138.

Halton, M. J. (2003). Benchmarking: Another attempt to introduce marketoriented policies into Irish second-level education? *Pedagogy, Culture and Society, 11*(3), 331–351.

Hatch, T. (2013). Beneath the surface of accountability: Answerability, responsibility and capacity-building in recent education reforms in Norway. *Journal of Educational Change, 14*(2), 113–138.

Hopkins, D., & Lagerweij, N. (1996). The school improvement knowledge base. In D. Reynolds, R. Bollen, C. Bert, D. Hopkins, L. Stoll, & N. Lagerweij (Eds.), *Making good schools. Linking school effectiveness and school improvement* (pp. 61–93). New York: Routledge.

Hopkins, D., & Reynolds, D. (2001). The past, present and future of school improvement: Towards the third age. *British Educational Research Journal, 24*(7), 459–475.

Kamens, D. H. (2013). Globalization and the emergence of an audit culture: PISA and the search for 'best practices' and magic bullets. In H. D. Meyer & A. Benavot (Eds.), *PISA, power and policy: The emergence of global educational governance* (pp. 117–140). London: Symposium Books.

Kauko, J., Rinne, R., & Takala, T. (Eds.). (2018). *Politics of quality in education. A comparative study of Brazil, China, and Russia*. Abingdon & New York: Routledge.

Kingdon, J. (1984). *Agendas, alternatives, and public policies*. Boston: Little, Brown, & Co.

Marcussen, M. (2004). Multilateral surveillance and the OECD: Playing the idea game. In K. Armingeon & M. Beyeler (Eds.), *The OECD and European welfare states* (pp. 13–31). Cheltenham, UK & Northampton, MA: Edward Elgar Publishing.

Martens, K. (2007). How to become an influential actor – The "comparative turn" in OECD education policy. In K. Martens, A. Rusconi, & K. Lutz (Eds.), *Transformations of the state and global governance* (pp. 40–56). London: Routledge.

Martens, K., & Jakobi, A. P. (2010). Introduction: The OECD as an actor in international politics. In K. Martens & A. P. Jakobi (Eds.), *Mechanisms of OECD governance: International incentives for national policy-making?* (pp. 1–25). Oxford: Oxford University Press.

Martens, K., & Niemann, D. (2013). When do numbers count? The differential impact of the PISA rating and ranking on education policy in Germany and the US. *German Politics, 22*(3), 314–332.

Martino, W., & Rezai-Rashti, G. (2013). "Gap talk" and the global rescaling of educational accountability in Canada. *Journal of Education Policy, 28*(5), 589–611.

McNamara, G., O'Hara, J., Boyle, R., & Sullivan, C. (2009). Developing a culture of evaluation in the Irish public sector: The case of education. *Evaluation, 15*(1), 101–112.

Møller, J., & Skedsmo, G. (2013). Modernising education: New Public Management reform in the Norwegian education system. *Journal of Educational Administration and History, 45*(4), 336–353.

Moos, L. (2014). Educational governance in Denmark. *Leadership and Policy in Schools, 13*(4), 424–443.

Morgan, C. (2009). *The OECD Programme for international student assessment: Unravelling a knowledge network*. Saarbrucken: VDM Verlag.

Niemann, D., & Martens, K. (2018). Soft governance by hard fact? The OECD as a knowledge broker in education policy. *Global Social Policy, 18*(3), 267–283.

Niemann, D., Martens, K., & Teltemann, J. (2017). PISA and its consequences: Shaping education policies through international comparisons. *European Journal of Education, 52*(2), 175–183. http://doi.org/10.1111/ejed.12220.

OECD. (2005). *School accountability and transparency of information.* Informal Meeting of OECD Education Ministers. Directorate for Education. EDU(2005)4.

OECD. (2007). *OECD's PISA survey shows some countries making significant gains in learning outcomes, press release.* Retrieved from http://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/oecdspisa-surveyshowssomecountriesmakingsignificantgainsinlearningoutcomes.htm.

OECD. (2008). *System accountability and the quality and equity of schooling outcomes. Recent work and issues for discussion*. Directorate for Education. EDU/ EDPC(2008)16.

OECD. (2010). *Reviews of national policies for education: Santa Catarina State, Brazil*. Paris: OECD Publishing.

OECD. (2011a). *Improving lower secondary schools in Norway 2011, reviews of national policies for education*. Paris: OECD Publishing.

OECD. (2011b). *PISA in focus, no. 9. School autonomy and accountability: Are they related to student performance?* Paris: OECD Publishing. https://doi.org/10.1787/5k9h362kcx9w-en.

OECD. (2013). *Synergies for better learning: An international perspective on evaluation and assessment*. Paris: OECD.

OECD. (2016). *Education in Colombia, reviews of national policies for education*. Paris: OECD Publishing.

OECD. (2017). *Education in Costa Rica, reviews of national policies for education*. Paris: OECD Publishing.

OECD. (2018). *Education for a bright future in Greece, reviews of national policies for education*. Paris: OECD Publishing.

Ougaard, M. (2010). The OECD's global role: Agenda-setting and policy diffusion. In K. Martens & A. P. Jakobi (Eds.), *Mechanisms of OECD governance:*

International incentives for national policy-making? (pp. 25–49). Oxford: Oxford University Press.

Pagani, F. (2002). Peer review as a tool for co-operation and change: An analysis of an OECD working method. *African Security Review, 11*(4), 15–24.

Pagès, M., & Prieto, M. (2019). *Re-contextualizing global education policies and reforming education through policy instruments: Evidence from Spain*. Working paper, Barcelona: GEPS.

Pal, L. A. (2009, July). *The OECD and global public management reform.* Paper presented at the 21st World Congress of the International Political Science Association, Santiago de Chile.

Papadopoulos, G. S. (1994). *Education 1960–1990: The OECD perspective* (OECD historical series). Paris: OECD.

Parcerisa, L., & Falabella, A. (2017). La consolidación del Estado evaluador a través de políticas de rendición de cuentas: Trayectoria, producción y tensiones en el sistema educativo [The consolidation of the evaluative state through accountability policies: Trajectory, enactment and tensions in the Chilean education system]. *Education Policy Analysis Archives, 25*(89), 1–27.

Popp, M. (2014). New culture, old system—Reactions to internationalization in Spanish education policy. In K. Martens & P. Knodel (Eds.), *Internationalization of education policy: A new constellation of statehood in education?* (pp. 163–178). London: Palgrave Macmillan.

Porter, T., & Webb, M. (2007, November). *The role of the OECD in the orchestration of global knowledge networks.* Paper presented at the International Studies Association Annual Meeting, Montreal.

Schleicher, A., & Zoido, P. (2016). The policies that shaped PISA, and the policies that PISA shaped. In K. Mundy, A. Green, & B. Lingard (Eds.), *Handbook of global education policy* (pp. 374–384). Hoboken: Wiley.

Schuller, T. (2005). Constructing international policy research: The role of CERI/OECD. *European Educational Research Journal, 4*(3), 170–180.

Sellar, S., & Lingard, B. (2013). The OECD and global governance in education. *Journal of Education Policy, 28*(5), 710–725.

Sellar, S., Thompson, G., & Rutkowski, D. (2017). *The global education race: Taking the measure of PISA and international testing*. Edmonton: Brush Education.

Sjøberg, S. (2016). OECD, PISA, and globalization: The influence of the international assessment regime. In C. H. Tienken & C. A. Mullen (Eds.), *Education policy perils: Tackling the tough issues*. New York: Routledge.

Tveit, S. (2009). Educational assessment in Norway – A time of change. In C. Wyatt-Smith & J. J. Cumming (Eds.), *Educational assessment in the 21st century: Connecting theory and practice* (pp. 227–243). Dordrecht: Springer Netherlands.

Verger, A., & Curran, M. (2014). New public management as a global education policy: Its adoption and re-contextualization in a Southern European setting. *Critical Studies in Education, 55*(3), 253–271.

Verger, A., Prieto, M., Pagès, M., & Villamor, P. (2018). Common standards, different stakes: A comparative and multi-scalar analysis of accountability reforms in the Spanish education context. *European Educational Research Journal*. https://doi.org/10.1177/1474904118785556.

Verger, A., Fontdevila, C., & Parcerisa, L. (2019). Reforming governance through policy instruments: How and to what extent standards, tests and accountability in education spread worldwide. *Discourse: Studies in the Cultural Politics of Education, 40*(2), 248–270.