This article was downloaded by: [Bibliothek PSYC & Ped Wet]

On: 14 June 2011

Access details: Access Details: [subscription number 931927228]

Publisher Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK

International Studies in Sociology of Education

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t716100714>

'Perform, measure accurately, optimise': on the constitution of (evidence-based) education policy

Mathias Decuypere^a; Maarten Simons^{ab}; Jan Masschelein^a

^a Laboratory for Education and Society, K.U. Leuven, Leuven, Belgium ^b Centre for Education Policy and Innovation, K.U. Leuven, Leuven, Belgium

Online publication date: 27 May 2011

To cite this Article Decuypere, Mathias , Simons, Maarten and Masschelein, Jan(2011) "'Perform, measure accurately, optimise': on the constitution of (evidence-based) education policy", *International Studies in Sociology of Education*, 21: 2, 115 — 135

To link to this Article: DOI: 10.1080/09620214.2011.575101

URL: <http://dx.doi.org/10.1080/09620214.2011.575101>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

‘Perform, measure accurately, optimise’: on the constitution of (evidence-based) education policy

Mathias Decuypere^{a*}, Maarten Simons^{a,b} and Jan Masschelein^a

^aLaboratory for Education and Society, K.U. Leuven, Leuven, Belgium; ^bCentre for Education Policy and Innovation, K.U. Leuven, Leuven, Belgium

(Received 17 November 2010; final version received 18 March 2011)

This article takes its point of departure in the current tendency of education policy to become more and more evidence-based. The use of statistics and numbers seems to be a prerequisite for conducting a policy that is both efficient and effective. The kind of knowledge thus produced is regarded as factual and scientific. This article tries to get a grasp on value-added modelling, a commonly used method supposed to produce such knowledge. Drawing on some conceptual underpinnings of actor-network theory, the article advances that such factual matters often take the form of matters of educational performance that are shaped and produced by means of calculative and inscription devices. However, the adagio that measures of performances should guide education policy is only one, albeit strong, point of view. Taking not only performances but also public issues into account could lead to what could be called a more ‘concern-oriented’ policy.

Keywords: evidence-based policy; value-added modelling; actor-network theory; educational performances

Introduction

Statisticians regularly bemoan what they perceive to be the lack of impact and appreciation of their discipline on the part of others. And it does seem to be true that we are faced with the paradox of more and more information being available, more and more complex problems to be solved – but, apparently, less and less direct appreciation of the role and power of statistical thinking. (Smith 2001, 17)

‘The times they are a-changin’. The citation above is a quote from Adrian Smith, one of the most influential advocates of the extension of the principles of evidence-based medicine to all areas of public policy, including education. By and large, Smith (1996) pleads for a society wherein evidence, open access

*Corresponding author. Email: mathias.decuypere@ped.kuleuven.be

to and communication of that evidence, reign. Stating that statistics hardly have any impact and are not appreciated enough seems, at least in education, at odds with current evolutions in policy making. Those evolutions have been documented over the last years in the research literature and seem to undertake two different, albeit compatible, directions: policy uses evidence (often obtained by means of statistics) to create and justify certain of its own actions, but it also furnishes evidence to inform and encourage actions of education institutes and citizens.

Hence, on the one hand, evidence is *used* in the shaping and legitimization of policy (e.g. Gorur 2011; Novoa and Yariv-Mashal 2003). Telling examples in this respect are comparative studies like the Program for International Student Assessment (PISA) or Trends in International Mathematics and Science Study (TIMSS). Purposes of such studies are the comparison and ranking of the educational system of different countries. The reach of the PISA-studies, for instance, is immense: more than 60 countries, representing 90% of the world economy, are participating in the OECD's PISA-studies nowadays. Contrary to national performance data, PISA allows international comparison of educational systems and performances (OECD n.d.). And as many have noted, this generation and use of comparative data is supposed to result in a new way of policy making: comparing data is a 'governing based on evidence' – a method focusing on evidence that legitimises political action and policy measures (e.g. Novoa and Yariv-Mashal 2003; Ozga 2009; Ozga and Lingard 2007, Rose 1999). On the other hand, policy also *provides* evidence that can be used not only by education institutes, but by citizens as well. The so-called 'league tables' illustrate this. In such rankings certain institutes (schools, hospitals etc.) are ordered from excellent to poor performing. This ordering is obviously not based on thin air, but on evidence, i.e. on performances on a number of deliberate indicators (Goldstein and Spiegelhalter 1996; OECD 2008; Teddlie and Reynolds 2000). The rankings thus produced can apply to primary schools, secondary schools or universities and schools for higher education (e.g. Dill and Soo 2005; Strand 1998). A lot of countries – Japan, Poland, the UK, the USA and so on – use such rankings. Motivations for the publishing of these tables are vast: one wants to give information to students and their families, if one has a ranking one can make a more informed school choice, education institutes will tend to ameliorate their position vis-à-vis other institutes etc. Another example of the provision of evidence are systematic reviews that aim to unveil 'what works' in schools and classrooms by meticulously gathering and synthesising (quantitative) data and results obtained in different academic, scientific publications (Slavin 2008a). The ultimate aim of these reviews is to gather results about effective classroom practice, on the one hand, and obtain a dissemination of these results and of 'what works' towards classrooms and schools themselves, on the other. This way, school leaders and teachers should be able to organise schooling and conduct teaching based on sound research evidence instead of (only) on bottom-up practice (Evans and Benefield 2001; MacLure

2005). In this respect, scholars have also been arguing the crucial importance of good/best practices. Anyhow, ‘evidence’ seems to be regarded from one particular standpoint, being evidence as most of the time numerical, but primarily comparative data that may assist in knowing and improving one’s (or others’) relative position in the educational field – that is, governing through evidence.

Despite these two different though intermingled and interconnected approaches towards evidence, one particular link between these two approaches is the method used. Whether policy provides or uses evidence, it is believed to be of the utmost importance that the indicators, rankings, measured performances etc. are as accurate as possible. This requirement is to be found almost everywhere amongst propagators of a more evidence-based educational system: it would be hardly justifiable to adapt characteristics of a certain system if one is unsure of the exactness of the evidence in question. Once this criterion of accuracy is met, evidence can lead to policy adaptations that create a better/higher performing school, classroom or educational system (Coe 1999; Slavin 2008a, 2008b). It is in this context that a particular method, i.e. value-added modelling, becomes salient. The value-added method is indeed precisely such a method that, according to the OECD, allows for ‘greater accuracy in measuring performance which then creates greater confidence in the interpretation of school performance measures’ (OECD 2008, 16). We will take value-added modelling – as proposed by the OECD – as a paradigmatic practice exemplifying the current tendency towards the use of ‘evidence’ (knowledge based on accurate measurements) in education in general and education policy in particular. As an ultimate and particular hallmark of how evidence is created, the analysis of the OECD practice of value-added modelling is a fruitful guide in debunking certain general mechanisms underlying evidence-based policy.

In what follows we will, first of all, and by means of further introduction, elaborate upon the value-added method – and in particular the OECD’s use of this method. More specifically, we will elaborate on the OECD’s perceived advantages of using this particular methodology in their construction of datasets (e.g. PISA-studies) that often take up the form of (and are interpreted as) evidence. In the following section we will, by means of some conceptual underpinnings derived from actor-network theory, try to shed some light on certain assumptions and working methods that go together with using value-added modelling as a particular *constructor* of evidence. In other words, we will investigate value-added modelling as a particular method for obtaining accurate results that seems to be more and more attractive. More particularly, the second section of the article will focus on certain significant *devices* at work when using value-added modelling that nevertheless rarely gain attention. In denominating educational realities ‘states of affairs’, we will furthermore pose that different sorts of policy – that is, different at the level of their ‘constitution’ – conceive of a particular state of affairs in different ways. We

will argue that evidence-based policy in general and value-added modelling as a concrete application in particular consider states of affairs as consisting of *matters of performance*: the educational reality as hidden for the layman's eye, but consisting of performances that can be measured and that hence result in an adagio of ever more thoroughly optimizing. We will argue that this particular *constitution* (that is, this particular set of beliefs and assumptions) of education policy is a highly modern one. This constitution defalcates its working methods, assumptions and the visibility of experts in trade for gains in efficiency, effectiveness and the construction of a spectacle of performances. Refusing the particular consequences of such seemingly neutral processes, we will propose in the final section a different, non-modern constitution that we will term a '*concern-oriented policy*' (departing from the notion 'matters of concern', e.g. Latour [2004a, 2004b, 2005]). In doing this and throughout this article, we position ourselves in a broader research field that takes a more critical stance towards evidence-based policy (e.g. Biesta 2007; Ozga 2008, 2009; Ozga and Lingard 2007; for an overview see Van de Perre, Simons, and Kelchtermans 2009).

The added value of value-added

Organisations like schools can be viewed from different perspectives. One can consider an organisation as open or closed, as mechanic or organic and so on. Value-added modelling thrives on a rational and open vision on organisations, where in a certain context inputs are converted through a process into outputs. Such a vision permits seeing educational activity as a process of adding value to the raw material of the input and this is exactly what value-added modelling is about: to measure how much value particular organisations (e.g. schools) add to students' achievements (Bennett 2005; Cave et al. 1997; Scheerens, Glas, and Thomas 2003). Crucial in value-added modelling is the notion of progress as increased performance. In order to measure this progress, and contrary to a lot of other measurement techniques, two moments of measurement are needed: a baseline measure at the beginning of a particular time period and an outcome measure at the end of the period. In accordance with guiding methodological principles, the measures should be valid, reliable and fit for value-added procedures: performance data need to be gained so that they mirror reality in the best possible way in order to guarantee and maximize potential school improvement (OECD 2008; Scheerens, Glas, and Thomas 2003). This accuracy is obtained by means of the double measurement: measuring at two points in time is carried out in order to avoid biased measures. For instance, if one measures a certain performance at only one particular point in time, other factors than just the net schools' contribution can play their role freely: some schools have a population of students that has on average a higher socio-economic status (SES) than other schools, which could be the cause of the higher performing of that particular school. Besides

SES, there are a lot of other possible biasing factors: gender, school size, intelligence etc. The value-added contribution of a school can thus be defined as:

The contribution of a school to students' progress towards stated or prescribed education objectives (e.g. cognitive achievement). The contribution is net of other factors that contribute to students' educational progress. (OECD 2008, 17)

and value-added modelling can be defined as:

A class of statistical models that estimate the contributions of schools to student progress in stated or prescribed education objectives (e.g. cognitive achievement) measured at at least two points in time. (OECD 2008, 17)

The rationale – the added value – of using value-added modelling is, according to the OECD (2008), threefold. First of all, it enables *school improvement initiatives*. A more accurate performance measure (that is, a double measure, which eliminates potential biasing factors) leads to a more accurate evaluation of the present situation. Indeed, improvement in a given activity requires an accurate evaluation of the current situation, which in turn requires an accurate measure of performance. Thus, improvement programmes are only effectively developed if the measurement of the actual performance is accurate (OECD 2008, 28–31). Value-added modelling is considered a valuable tool in this respect, since it avoids raw test scores that are the result of measuring at one point in time. Hence the prime argument for using value added modelling: the provision of more accurate measurements than other and previous methods that investigate school performances:

[V]alue-added modelling provides more accurate estimates of school performance than do the comparisons of raw test scores or cross-sectional contextualised attainment models ... (OECD 2008, 11)

Improvement initiatives that become possible due to value-added modelling are, for instance, identifying those areas that are adding the most value and those that are in need of improvement; suggesting which strategies do and don't work; creating projections that can assist in planning, resource allocation and decision-making; setting performance targets; etc. A second rationale for the use of value-added modelling is that it is beneficiary for *systems of school accountability*. Here also, key issue is whether the assessment of the performance in question is accurate and, thus, fair. With accurate measures, one can hold accountable those entities that are responsible for the performance of a certain practice or output. Value-added modelling, according to the OECD, avoids making precipitate conclusions based on results that could be clouded by bias in measures, due to bias in gender, SES etc. This is the reason why the method is considered to be more honest, or fair, towards all parties involved, since only what is learnt in school is measured – not what was already known before school entrance. This fairness allows holding parties accountable for

particular performances: since the measures are accurate, one can claim responsibility for those performances. Consequentially, the general public, i.e. the taxpayer, might be informed whether tax money is spent efficiently, schools might be evaluated more effectively etc. The last rationale, closely linked to previous one, concerns *school choice*. Choosing a school benefits from value-added modelling as well. Indeed, when published, parents are informed of the (value-added) performance of different schools. This can aid their decision-making in choosing a particular school for their child. Benefits are, amongst others: that one can choose a school better suited to what is needed; that resources can be allocated more correctly; and so on. In this respect as well (as with the other two rationales), it is of particular importance that measurements are accurate. Otherwise, disadvantageous consequences, such as a misinforming of families' choices and a misallocating of resources, might come into play. Hence, also regarding this third rationale, what is at stake seems to be the crucial importance of accurate measurement of performances:

The greater accuracy inherent in value-added modelling creates greater confidence in the use of performance measures to further the three policy objectives [school improvement, school accountability, school choice] outlined above. (OECD 2008, 14)

What thus seems to be of central interest in each of the three rationales when using value-added modelling, is a greater accuracy compared with other and previous methods of measurement. This accuracy makes the central methodological criterion of reliability safeguarded and guaranteed: the more accurate one measures, the more one is certain that the resulting measures are reliable. This reliability is of central importance if one wants the stakeholders of a school to have confidence in the effectiveness of a certain school system (OECD 2008, 193). This confidence, on its turn, is considered to act as impetus for the establishment of incentives of different sorts. Highly reliable performance measures alone might do the trick – or rather, the more reliable the measure, the more compelling the incentives:

The construction and publication of performance measures *alone* might provide implicit and indirect incentives to school principals and teachers. ... The government, by disclosing information about school performance and perhaps connecting rewards to performance, might induce teachers and administrators to respond by increasing their efforts to increase performance ... (OECD 2008, 36–37, emphasis added)

Thus, disclosing performance measures seems to deliver incentives of some sort. These incentives are dividable into two different categories. First of all, there seems to be a conviction on part of the OECD that if one connects 'rewards to performances', teachers and administrators will be inclined to increase efforts, that is, the reward as *explicit and direct* incentive for school

improvement (or, rather, performance improvement). Furthermore, as the first part of the quotation and more particularly the use of a word as ‘alone’ suggests, incentives can also be rather *implicit and indirect*: constructing and publishing alone apparently suffices for causing attempts towards increase of performances – as if, so to speak, the facts speak for themselves.

This chain ‘accuracy – reliability/confidence – incentives’ is, however, not that evident from the outset. For instance, why and for whom do ‘facts speak for themselves’? How come the publishing of measurements *alone* already provides incentives? Which processes are at work in order to come to this rather peculiar chain? In what follows, we will try to answer these questions and argue that in order for this chain to be established and operationalised, what we call *socio-technical devices* are needed. Socio-technical devices incorporate and perform specific operations (that is, the technical aspect of the device), on the one hand, but also need a particular self-understanding of the so-called stakeholders of a school (that is, the social aspect of the device), on the other.

Calculative and inscription devices

In the description of value-added modelling in the previous paragraphs, we came to the observation that this method not only has some new characteristics (e.g. greater accuracy and hence greater reliability), but that users of either the method or of its results share some conviction that the measures value-added modelling produces ‘speak for themselves’. Hence, value-added modelling not only has a technical-methodological component, but some social-normative component as well. In line with actor-network theory, which took as one of its early points of interest the study of the creation and genesis of scientific facts (Latour 1987; Latour and Woolgar 1986), we contend here that the creation of performance measures by means of value-added modelling and, consequently, of the treatment of performance measures as ‘facts/evidence’ that speak for themselves, is dependent on an assemblage of social and technical elements. More specifically, we will consider value-added modelling not only as a method, but as a socio-technical device as well – this device neither consisting of technical elements alone, nor fully determined by mere social processes. Considering value-added modelling as a device makes it possible to look at *what it does* and at which different operations it performs – briefly, to look at its distributed agency. We will explore this distributed agency from four different angles: (1) the assumptions of improvement, accountability and choice, (2) the creation of a space of equivalence, (3) the creation of a field of identity and difference and (4) responsabilisation.

(1) *The assumptions of improvement, accountability and choice*

As was made clear in the preceding section, the OECD’s orientation towards current educational states of affairs consists in a conviction that every practice

can be improved and that this improvement can only take place if the methods of measurement are as accurate as possible. This confidence in a potential improvement for every educational situation imaginable is, it is assumed, also present in other stakeholders of the policy process. Teachers, principals, governments – all are (considered to be) eager to increase their scores on what is measured, for instance people involved in the shaping of policy:

Once value-added information is obtained, decision-makers can better analyse how to adjust resources and enact the appropriate school processes to improve student performance. The inclusion of value-added information allows organisational learning of what best contributes to performance improvements. (OECD 2008, 28)

This radical assumption of improvement (and consequential assumptions of accountability and choice) is even to be found at the utmost beginning of the book, being the first sentence of the foreword:

As OECD countries seek to improve their education systems, a growing emphasis is being placed upon measures of school performance as they are central to school improvement efforts, systems of school accountability and school choice, and broader education policies. (OECD 2008, 3)

Improvement is framed here in a particular way. The criterion for improvement is a technical one, that is, something improves when it obtains a better performance (more efficient, more effective or a more optimal input-output balance) (Lyotard 1979). This is, in a nutshell, the first operation needed in order to speak of facts as speaking for themselves – this assumption of improvement is the first operation, or act of agency, of value-added modelling as a socio-technical device: the underlying assumption that every educational practice can be framed in view of its performance and that increased performance is the criterion for improvement (and consequential accountability and choice).

(2) *The creation of a space of equivalence*

A second important aspect of value-added modelling as socio-technical device is its enactment of equivalence. Key to understanding the self-evidence of the chain discussed above is the OECD's emphasis on what it terms as 'comparability'. This comparability between schools, universities, national education systems etc. is, however, not achieved automatically and has to be constructed (see also Gorur 2011). It is the *construction* of measures that 'might provide incentives' (OECD 2008, 36) and this construction is necessary in order to obtain the required indicators:

In most participating countries the rationale for implementing a value-added system ... is to focus the attention of school leaders, teachers and students on

improving on those measures and student learning in the corresponding academic disciplines. ... Typically, student *scores are transformed* or summarised into performance indicators that inform the decision-making process. (OECD 2008, 119, emphasis added)

This construction or transformation of measures is necessary to install a 'space of equivalence' in which performance measures are brought onto one single scale, facilitating the perceiving of differences and consequential comparability. Thence the constitution of a commensurate space of international comparison; a space of uniform measurement (Rizvi and Lingard 2009, 448). This space is characterised by common mental representations (e.g. schools can be improved), a common language (e.g. talking in terms of 'added value' and 'value added') and common measurements (e.g. the usage of the same categories and procedures of coding for all countries, as is effectuated in for instance the PISA studies) (Desrosières 2002; see also Ozga 2008). Once a space of equivalence is constituted, the major challenge turns out to be how to make those common measures as accurate as possible – hence making this very space even more taken-for-granted. Equivalence is constructed by means of the (accurate) output numbers of the method, is reinforced by focusing on accuracy and is ultimately leading to a common, taken-for-granted scale where schools and other institutions can be compared according to their performance. In other words, output numbers are reflections of a school's *relative* performance:

Once the [value-added] model is applied to the data the output is a set of numbers, one for each school. ... The number attached to a particular school is provisionally interpreted as a measure of the school's relative performance... (OECD 2008, 106)

A space of equivalence thus makes it possible for schools to know their performance in relation to other schools. Additionally, this space is at once a particular space of improvement (as well as of choice and accountability), it frames improvement as increased performance, locates responsibilities and creates objects of choice and accountability. Hence, the space of equivalence/improvement is not to be considered as out there and existing 'as such', but is *constituted* and this constitution includes several social operations.

(3) *The creation of a field of identity and difference*

The inauguration of a space of equivalence gives way to another operation performed by value-added modelling as socio-technical device, viz. the instalment of a field of identity and difference. This instalment of differences is not something that is given in the natural order of schools (though made equivalent): in order to have *differences* between schools, schools have to be made *differing* in view of performance output and added value. This is effectuated by means of *calculation*, a technique performed by *calculative devices*

that reside in the method of value-added modelling. Drawing loosely on the work (and words) of Michel Callon and colleagues (Callon and Muniesa 2003; Callon, Méadel, and Rabeharisoa 2002; Muniesa, Millo, and Callon 2007), we perceive calculation as the perceiving and grading of differences in a space of equivalence, which lead to estimating courses of action associated with the perceived differences. Calculative devices, then, are devices that aim at rendering things calculable, or more precisely, at enacting particular versions of what it is to be 'calculable'. In the previous paragraphs, we showed that those differences can only come to the fore when there is a common horizon and common measurements – in short, when there is a space of equivalence in which to create and perceive the differences. But what, then, are the means by which those differences are perceived and graded? In the document on value-added modelling, the OECD discusses a couple of instruments that can be considered as calculative. Rankings are one example in this respect (OECD 2008, esp. chapter 2). The calculative nature of a ranking is situated in its very attempt to differentiate between the good and the bad, the high- and the low-performing. Furthermore, a ranking is a monitoring instrument: it tries to detect low-performing schools that have to improve their performance in order to keep up with other schools (see above). Value-added modelling, as a socio-technical device including several calculating instruments, inscribes a field of objectivity that makes it possible to experience identity and difference in terms of performance levels. In other words, it makes possible processes of identification and differentiation based on performance levels. Those experiences of identity and difference in terms of performance (that is, 'my identity in terms of my performance' in comparison to 'performances of others') actually induce a rationale of improvement for schools, a rationale of choice by parents and a rationale of accountability for governing agencies.

Besides these calculative instruments – which can be concrete objects, concepts, tools, methods etc. – other devices, i.e. *inscription devices*, are needed in order to sustain identity and differences. Inscription is 'a general term that refers to all the types of transformations through which an entity becomes materialized into a sign, an archive, a document, a piece of paper, a trace' (Latour 1999, 306). Stated otherwise, inscription devices effectuate a process by which certain aspects of reality are transformed into material entities/accounts – e.g. figures, diagrams and texts (Latour 1987, 1999; Law 2004; Rose and Miller 1992). Calculative devices, for instance, inscribe very particular features of a particular situation into a certain state of affairs: only those aspects that are being measured are inscribed, for instance as a number in a table, as a particular position in a ranking, as a concrete position in a caterpillar plot, as a benchmark to reach etc. The various understandings such indicators and benchmarks perform are, amongst others, to understand oneself as in need to raise the effectiveness of schools and policy measures, to make more informed choices, to be able to compare between periods of time or between different education

systems, which in turn might lead to rankings of schools and major comparative studies like PISA. The interrelatedness of all such calculative devices is remarkable and can be understood by considering calculation and inscription as two sides of the same coin: in order to calculate, one needs inscription.

These processes of calculation and inscription create comparability on two levels: first of all because of the 'comfort of numbers' (a number, a table ... is more easy to compare than two detailed inspection reports, for instance), secondly because data of a certain school can be transferred on the spot, be sent by email to the other side of the world, be printed in a publication etc. Thus, by means of calculative and inscription devices, a state of affairs is becoming susceptible for evidence-based measures and evidence-based government. Furthermore, those measures can start to be incentives of some sort themselves: accurate measures perform a specific logic of improvement, they tend to point towards higher numbers, which are then considered to be *better* numbers. Measures thus give incentives by means of a *logic of amelioration*, being (1) that education systems have to perform, (2) that these performances have to be measured accurately and (3) that, whatever the outcome of those measures may be, one always has to try to improve the particular performance at a certain point in time. The objectivity that is inscribed and that opens up a space to experience identity and differences at once translates the logic of improvement as the amelioration of objective figures. This logic of amelioration makes particularly clear that inscription is no mere neutral process or product. Rather on the contrary, it *produces* objectivity, it is a technique of thought that makes possible the extension of authority over that which it seems to depict (Rose 1999, 37). Inscription devices make states of affairs stable, comparable and governable.

(4) *Responsabilisation*

The actual actions (of improvement, choice, accountability) are only possible when actors come to understand themselves (and/or others) as responsible for their (improvement of) performance. Here, it is important to stress that the space of equivalence/improvement and generated identities and differences install at once specific loci of responsibility for output performance. The measurement of the added value of a school, for instance, is at once framing the school as an agent responsible for the performance level (in view of improvement, choice and accountability). Again, the device called 'value-added modelling' differentiates on the one hand, and triggers specific actions based on the (scientifically) made differentiation (and, hence, identities) on the other. This triggering of actions is, however, not that self-evident. In order to trigger actions, it is first of all necessary for people, policy etc. to come to understand oneself, the education system and so on, as measurable and in need of improvement. Value-added modelling tends to make actors understand themselves in this process as the locus of responsibility, it installs the previously discussed adagio of

improvement (as amelioration) as the (evident) horizon of what to do and how to think, for instance the premise that statistical analyses are sufficient in order to motivate policy adjustments or the assumption that regardless of the context of the country in which value-added modelling was applied, there will always be an inclination towards improvement:

Statistical analyses of the relations between school inputs and indicators of school performance can suggest which strategies are and are not working, leading to policy adjustments and the reallocation of resources. (OECD 2008, 12–13)

Explicitly identifying how value-added scores will be interpreted and used to trigger specific actions increases the level of transparency and internal efficiency. Stakeholders need to know these actions in order to have confidence in the system and also to design appropriate measures to lift performance. (OECD 2008, 207)

This paragraph offered insight in the socio-technical dimensions of value-added modelling as a method: value-added modelling is not just a pair of glasses enabling to see (or, mirror) reality ‘for real’, but includes some devices and technologies that construct a particular state of affairs through specific measures. The current focus on more accuracy is actually ‘black-boxing’ particular construction mechanisms and reinforcing the constructed common horizon and space of equivalence. Value-added modelling neither consists of technical elements alone, nor is it fully determined by mere social processes. Instead, agency (improvement, accountability, choice) is distributed amongst the materiality of a particular instrument and the person (or institution) who handles the instrument. Although it is assumed that the final outputs or rankings ‘speak for themselves’ and are an incentive for improvement, choice or accountability, it is actually these devices and technologies that ‘make’ them speak for themselves, distribute the required agency and generate the experience of identity, difference and responsibility. In the next paragraph, we will try to unravel some commonly held beliefs and fundamental assumptions (that is, a particular constitution) coupled to this value-added modelling of the current state of affairs in education.

A highly modern constitution

In order to get further grasp on the OECD’s stance towards the state of affairs in the various education systems of today, we will draw on Bruno Latour’s theming of the use of politics and science in present societies. According to Latour, the constitution of a great many Western societies can be typified as ‘modern’. This modern constitution holds as one of its main tenets a divide between Science, on the one hand, and society, on the other. By this, Latour means that modern societies are characterised by a break, and according

differences in appreciation, between what is considered to be ‘really’ scientific and what is considered to be ‘mere’ social. This distinction is achieved by means of two ‘collectors’ (proverbial containers with distinct ontological contents): one collector for human beings in which social events occur (forming bonds, doing politics, talking about and creating values), another for natural objects of all kinds (objects, facts, laws of nature etc.). The bifurcation between nature and society, mirrored by the two collectors, is furthermore exemplified by the content and the concomitant valorisation of both collectors. The nature-collector is stuffed with *matters of fact*: indisputable entities that take in an unproblematic place in the coexistence of human beings and things. Matters of fact can hence be regarded as what we commonly call ‘facts’, ‘objects’ or ‘nature’ or what we refer to as ‘objective’ (Latour 1999, 2004a). The society-collector, by contrast, contains what could be called *matters of value*: disputable human opinions, struggles, argumentations, interpretations, values – those things which are in modern terms depicted as ‘social’ or ‘sociality’ and which are, hence, subjective, human characteristics.

In this paragraph, we will proceed by showing that value-added methods and their corresponding use are fitting this modern constitution in a very peculiar way. More specifically, we will argue that the constitution of value-added modelling is not just modern, but rather what we will call ‘highly’ modern. As was already made clear above, the OECD’s prime rationale for using value-added modelling is the premise that (greater) accuracy leads to better policy: ‘value-added analysis can be used to advance a number of policy objectives’ (OECD 2008, 175). Putting central accurate measurements implies a strong valuation of scientific measurement: the more accurate the measurement, the greater the probability what one measures is actually what one wants to measure, the closer one gets to the ‘real’ matters of fact. Or, to put this in the terms introduced above, the better one knows (or observes, or registers) the matter of fact at hand (‘evidence’), the more policy can undertake appropriate action in the society-collector. This is, in a nutshell, what evidence-based policy is all about: the implementation of scientific knowledge in the daily practice of education and policy in the various ways outlined in the introduction (Oakley 2002; Slavin 2008a, 2008b). There thus seem to be, indeed, two collectors and a corresponding division between matters of fact (as ‘evidence’) and matters of value (as ‘policy’ based on that evidence), noticeable in evidence-based policy. However, the term ‘matters of fact’ is maybe not fit enough and too broad to capture what happens in evidence-based policy specifically. As most of the extracts above suggest, the matters of fact in question almost always deal with (indicators of) performances and/or performance measures. Value-added modelling is precisely a method that enables conducting performance measures:

[I]t is important to note that value-added models overcome many of the distortions associated with other school performance measures. (OECD 2008, 72)

Thus, the matters of fact the OECD is talking about are mainly ‘facts about performances’ or ‘facts about performance measures’. This focus on performances in evidence-based policy is a widely recognised phenomenon: research has amply shown that quantitative information in evidence-based policy often takes up the form of information about performances (e.g. Grek 2009; Ozga 2009; Simons 2007; Van de Perre, Simons, and Kelchtermans 2009). However, indicators that are the resultant (output) of value-added modelling immediately contain certain *values* as well. Value-added measurements are, in other words, evaluative, since these measures contain *in themselves* already certain values, which one can hardly argue about (cf. the logic of amelioration mentioned above). Comparative information always induces a kind of drive to do better and to perform higher (Ball 2003; Rose 1999). Hence, one could state that value-added measures possess an *immanent normativity* (Berns 2009), including at once an objective statement about a given state of affairs, on the one hand, and an evaluative statement opening up a space of reconstruction (or improvement as amelioration), on the other – these evaluative statements being the normative component, the assumption that every education system can/should be improved (either against earlier performances or against equivalent actors).

This analysis enables to specify and retranslate the modern distinction between (matters of) fact and (matters of) value, as far as evidence-based policy and value-added modelling is concerned, as a merged field or space of (*matters of*) performance and (*matters of*) optimization (*choice/accountability*). Indeed, in evidence-based policy the measured performances act as and are considered to be facts; the value linked hereto is optimization (and the criterion for optimization is contained in the measure and concomitant comparison itself). This highly modern constitution, as we would like to call the constitution that enables the existence and sustainment of this merged field, gives way to particular roles for researchers and policy. Researchers, on the one hand, are no longer visible as highly competent, authoritative figures as is the case in a modern constitution. Rather, they seem to have a less visible position as the authority lays in the performance measure itself. A ranking, for instance, is assumed to ‘speak for itself’ in the sense of opening up a space of equivalence and a space where improvement is framed as increased ‘performance’ or ‘added value’. Policy, on the other hand, deals with the optimization of education and the taking of measures that try to attempt to add as much value as possible. Policy does no longer need to make judgements about what to do with the measures, but can immediately start to optimize after the performances are measured. Summarizing thus far, we approached evidence-based policy and value-added modelling in terms of a highly modern constitution. The new concepts proposed here, namely *matters of performance* and *matters of optimization*, are specific applications of this analysis to the field of evidence-based policy. In this field, reality is inscribed in terms of performances, the

criterion for optimization is located within reality itself, and leading to the assumption that policy's and other stakeholders' task is the improvement of performances whatsoever.

In this section we undertook an attempt to get a grasp of the precise constitution of evidence-based policy (and value-added modelling as a particular method) and how this constitution perceives current educational states of affairs. In doing this, we argued that evidence-based policy regards those states of affairs as open to objective assessments in terms of what is and what is not effective (i.e. performing well), of what is to a greater or lesser extent efficient (i.e. performing well in relation to the cost it takes to perform that well) – and that the according performance measurements contain an immanent normativity that locates the criterion for potential improvement within the measured state of affairs. This highly modern constitution tends to inaugurate a process of invisibility of expertise and experts. Because of the evaluative character of the value-added measurement, and what the proponents refer to as its potential to give 'incentives', the process of construction of those measurements tends to disappear. Which scientists are responsible for the construction of the PISA-questions? Who is behind TIMSS or PIRLS? Sure, we can depict some organisation, but the scientists behind them are hardly known by anyone. After all, don't the facts speak for themselves? The invisibility of experts stands in sharp contrast with the visibility of performances, which tend to take up the form of a 'spectacle' of performances in which each and all is oriented towards those institutes or other agencies that perform the best. What is ultimately installed within this spectacle is the longing to become an image of good performance, of taking a position in the centre of the spectacle and being watched constantly and rigorously, of one's performance acting as an incentive itself to the mass watching the few (Simons 2007, 543–44). The permanent focus on performances, the profound competition between institutes, the unconditional confidence that that what is measured needs to be optimised, all are serious preoccupations that have been criticised over the last few years (e.g. Ball 2003; Ozga 2009; Prickett 2002). Facing these challenges, perhaps we could and should explore a different constitution for education policy than the highly modern one, where facts speak for themselves, and elaborate upon a concrete (new) proposition that focuses not only on performances, competition, accuracy, incentives and the like. In order to do this, we will once more draw on Latour and introduce his specific understanding of what is 'public'.

Discussion and conclusion: having the answer versus being concerned

According to Latour (2005), the public has to be understood in a pragmatist sense. In line with Dewey, something becomes public when there is a certain issue that cannot be solved with available scientific expertise and knowledge. Dewey (1954) depicts the public as:

All those that are affected by the indirect consequences of transactions, to such an extent that it is deemed necessary to have those consequences systematically cared for. ... The essence of the consequences which call a public into being is the fact that they expand beyond those directly engaged in producing them. (15–16)

With such a pragmatist stance in mind, Latour (2007b) considers the public as those who are focused on such consequences and who produce the instruments and equipment necessary to let them appear and to deal with them. Close to the thoughts of Marres (2005), Latour states that a public is constituted by those who are gathered around an issue that *cannot* be allotted to one particular expert agency or the other, for they too don't have the answers. These issues can be regarded as *matters of concern*: an expression invented to contrast with matters of fact, and pointing to contested states of affairs. The highly modern distinction between matters of performance – matters of optimization we elaborated upon before, can with this new concept of what is 'public' be retranslated in a distinction between matters of concern – public: a public is gathered around (but not opposed to) a certain matter of concern. This means a matter of concern is created, surrounded and dealt with by a public, and in this public no one can define or claim a certain state of affairs and the particularities that circulate in it beforehand – this is precisely what makes the public a public (Callon, Lascoumes, and Barthe 2009; Latour 2004, 2005). We think it can be fruitful not to take a highly modern but a *non-modern* approach as point of departure. In such a non-modern constitution, it is impossible to state that 'facts speak for themselves'. Rather, this issue should be considered more as a *controversy* than as a fact 'as such', a controversy where different parties have different approaches and different propositions for dealing with the state of affairs at hand (Venturini 2010). This diversity in approaches and solutions is, however, not a plea for some sort of 'anything goes'. On the contrary, at some point in time, some kind of closure needs to be achieved (Latour 2004). In order for closure to be achieved, what is of importance is that different *spokespersons* speak and that those spokespersons raise different approaches towards the matter(s) of concern at hand – not for *themselves*, but for that which or whom they represent (on this notion of spokesperson, see, e.g., Callon [1986]). Tentatively, we can put forward the issue/controversy 'what constitutes good education' as a possible matter of concern. Indeed, along with the (highly) modern constitution, this issue can end up being treated in terms of performances and optimizations and the corresponding inauguration of spaces of equivalence where comparability and responsabilisation are commonplace (previous sections). However, it can appear in terms of matters of concern around which a public is gathered as well (what follows). A *concern-oriented policy*, as we would like to term this alternative viewpoint, would thus refuse to reduce the debate regarding 'education' to an allotted matter (that is, 'known' and, hence, in hands of a small group of experts), but would treat such a debate as an issue or a matter

of concern, which has the potential to generate a concerned public centred around this issue.

A concern-oriented policy is, in other words, an explicit attempt to rethink policy in a non-modern way – an attempt to open up and to keep open ‘the possibility that things might be done differently; the permanent possibility of alternative configurations’ (Mol 2002, 164).¹ However, this attempt to open up possibilities as we suggest it here does not depart from the centrality of performances, but rather from central questions as ‘What reality should we live with?’ (165) or ‘What constitutes good education?’ and raises possibilities for different spokespersons to enter the public concerning the issue at hand. To paraphrase Isabelle Stengers (2005), we live in a reality unknown, constituted and characterised by multiplicity and divergence – a *cosmos* where we cannot say in advance what is good (and what is not), what has to be done (and what not). This implies the abandonment of any space of equivalence whatsoever (‘for equivalence implies a common measure and thus an interchangeability of positions’ [1995]) in favour of a *space of equalization*, a space where the issue of the ‘good’ stands central and where differences are not flattened out right away (Mol 2002; Stengers 2005, 996). In such (cosmic) policy-making, some spokespersons can speak on behalf of the performance of the system. Some might speak on behalf of the labour market and plea for good education as education that provides the right qualifications and competencies the economy needs. Even others might speak on behalf of education to have a right of its own where children have a right to be ‘just’ children and where the focus does not have to be on what society needs. *It does not matter*, as long as all those groups compose together a public of spokespersons gathered around the issue ‘education’. In a similar vein, methods as value-added modelling can be helpful in trying to get to an answer. Calculative devices as rankings and benchmarks, examples of good performance, systematic reviews and so forth can be fruitful as well, on the condition that they are used by spokespersons who are on the same level as the other spokespersons in the public forums – as ‘spokespersons’, since ‘no calculative device is a substitute for political decisions’ (Latour 2007a, 5). Hence, in a concern-oriented education policy, as we envisage it here, the attempts, propositions, answers, solutions and measures will always be those of the public – and never *a priori* those of a certain accurate measurement produced by a scientific authority. Consequentially, spokespersons will have to make use of modes of assemblage that allow spokespersons to act as spokespersons *for something* (be it the issue ‘education’ or specificities concerning this issue, e.g. performances, well-being, teachers, government etc.). For instance, one could prevent the inclination to let numbers (‘facts’/‘evidence’) speak anonymously for themselves as is done in a highly modern constitution by having them *represented* by a spokesperson of the group of researchers that *produced* them by means of calculative devices. As spokesperson they would not be able to claim that those numbers are the only possible factual representation of the educational reality, but, on the contrary, bring

an argument based on such numbers into the public debate (or assemblage). The difference with some highly modern form of thinking is that researchers/experts as spokespersons thus become part of the argument and discussion and that numbers/performances do not have the chance to inscribe themselves into reality as soon as they are produced: letting a spokesperson make an argument based on numbers is something totally different than letting numbers speak for themselves. Additionally, in order for them to be able to represent someone or something, spokespersons will have to be gathered into the space of equalization: they will have to be *assembled* around a particular issue (whether that is by means of democracy, parliaments, public forums etc.):

Thanks to the notion of spokesperson, a process of assembling can now begin, one that no longer divides up the types of representatives in advance according to whether they demonstrate what things are or declare what humans want. In the single ... forum, each of the interested parties can, at a minimum, agree to consider the other as a spokesperson, without finding it relevant to decide whether the other represents humans ... the ... economy, nongovernmental organisations, or elected governments. (Latour, 2004, 65)

Again, this assemblage will have to be performed in order to bring spokespersons together and to create a public, but maybe this search for public modes of assemblages is a fruitful future task to undertake – for policy, teachers, students and researchers the like. This assumes, however, that we embrace the entrance of a non-modern constitution – being concerned with equalisation, instead of looking for equivalences.

Note

1. On this point, it would be interesting to trace how other, alternative conceptions of (evidence-based) policy could relate to the particular take on (evidence-based) policy we advocate here, for instance a more holistic approach towards ‘good practices’ (see, e.g., Hargreaves 2003).

Notes on contributors

Mathias Decuyper is a PhD researcher at the Laboratory for Education and Society, Katholieke Universiteit Leuven, Belgium. His primary research interests are the changing nature of education policy, socio-technical approaches, the adoption of new media in education and higher education in particular.

Maarten Simons is a Professor in the Center for Educational Policy and Innovation and the Laboratory for Education and Society, Katholieke Universiteit Leuven, Belgium. His research interests are in educational policy and political philosophy with special attention to governmentality and schooling, and the ‘public’ character of education.

Jan Masschelein is a Professor in the Laboratory for Education and Society, Katholieke Universiteit Leuven, Belgium. His primary areas of scholarship are educational

theory, political philosophy, critical theory, studies of governmentality, and the 'public' character of education.

References

- Ball, S.J. 2003. The teacher's soul and the terrors of performativity. *Journal of Education Policy* 18, no. 2: 215–28.
- Bennett, N. 2005. Power, structure and culture: An organizational view of school effectiveness and school improvement. In *School effectiveness and school improvement. Alternative perspectives*, ed. A. Harris and N. Bennett, 98–122. London: Continuum.
- Berns, T. 2009. *Gouverner sans gouverner: Une archéologie politique de la statistique* [Governing without governing. A political archaeology of statistics]. Paris: PUF.
- Biesta, G. 2007. Why 'what works' won't work: Evidence-based practice and the democratic deficit in educational research. *Educational Theory* 57, no. 1: 1–22.
- Callon, M. 1986. Some elements of a sociology of translation: Domestication of the scallops and the fishermen of St Briec Bay. In *Power, action, and belief: A new sociology of knowledge*, ed. J. Law, 196–233. London: Routledge and Kegan Paul.
- Callon, M., and F. Muniesa. 2003. Les marches économiques comme dispositifs collectifs de calcul [Economic markets as calculative collective devices]. *Réseaux* 21, no. 122: 189–233.
- Callon, M., P. Lascoumes, and Y. Barthe. 2009. *Acting in an uncertain world: An essay on technical democracy*. Trans. G. Burchell. Cambridge, MA and London: MIT Press.
- Callon, M., C. Méadel, and V. Rabeharisoa. 2002. The economy of qualities. *Economy and Society* 31, no. 2: 194–217.
- Cave, M., S. Hanney, M. Henkel, and M. Kogan. 1997. *The use of performance indicators in higher education*. London and Bristol: Jessica Kingsley.
- Coe, R. 1999. A manifesto for evidence-based education. Centre for Evaluation and Monitoring. <http://www.cemcentre.org/renderpage.asp?linkID=30317000>.
- Desrosières, A. 2002. *The politics of large numbers: A history of statistical reasoning*. Cambridge, MA and London: Harvard University Press.
- Dewey, J. 1954. *The public and its problems*. New York: Holt.
- Dill, D.D., and M. Soo. 2005. Academic quality, league tables and public policy: A cross-national analysis of university ranking systems. *Higher Education* 49, 495–533.
- Evans, J., and P. Benefield. 2001. Systematic reviews of educational research: Does the model fit? *British Educational Research Journal* 27, no. 5: 527–41.
- Goldstein, H., and D.J. Spiegelhalter. 1996. League tables and their limitations: Statistical issues in comparisons of institutional performance. *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 159, no. 3: 385–443.
- Gorur, R. 2011. ANT on the PISA trail: Following the statistical pursuit of certainty. *Educational Philosophy and Theory*. <http://onlinelibrary.wiley.com/doi/10.1111/j.1469-5812.2009.00612.x/abstract>.
- Grek, S. 2009. Governing by numbers: The PISA 'effect' in Europe. *Journal of Education Policy* 24, no. 1: 23–37.
- Hargreaves, D.H. 2003. *Education epidemic: Transforming secondary schools through innovation networks*. London: DEMOS.
- Latour, B. 1987. *Science in action*. Cambridge, MA: Harvard University Press.
- Latour, B. 1999. *Pandora's hope: Essays on the reality of science studies*. Cambridge, MA: Harvard University Press.

- Latour, B. 2004a. *Politics of nature: How to bring the sciences into democracy*. Trans. C. Porter. Cambridge, MA: Harvard University Press.
- Latour, B. 2004b. Why has critique run out of steam? From matters of fact to matters of concern. *Critical Inquiry* 30, no. 2: 225–48.
- Latour, B. 2005. From Realpolitik to Dingpolitik or how to make things public. In *Making things public. Atmospheres of democracy*, ed. B. Latour and P. Weibel, 14–43. Cambridge, MA: MIT Press.
- Latour, B. 2007a. How to think like a state. Paper presented at the occasion of the anniversary of the WRR, November 22, in Den Haag, The Netherlands.
- Latour, B. 2007b. Turning around politics: A note on Gerard de Vries' paper. *Social Studies of Science* 37, no. 5: 811–20.
- Latour, B., and S. Woolgar. 1986. *Laboratory life: The construction of scientific facts*. 2nd ed. Princeton, NJ: Princeton University Press.
- Law, J. 2004. *After method: Mess in social science research*. London: Routledge.
- Liotard, J.-F. 1979. *La condition postmoderne: Rapport sur le savoir* [The postmodern condition: A report on knowledge]. Paris: Minuit.
- MacLure, M. 2005. 'Clarity bordering on stupidity': Where's the quality in systematic review? *Journal of Education Policy* 20, no. 4: 393–416.
- Marres, N. 2005. Issues spark a public into being: A key but often forgotten point of the Lippmann-Dewey debate. In *Making things public: Atmospheres of democracy*, ed. B. Latour and P. Weibel, 208–17. Cambridge, MA: MIT Press.
- Mol, A. 2002. *The body multiple: Ontology in medical practice*. Durham, NC: Duke University Press.
- Muniesa, F., Y. Millo, and M. Callon. 2007. An introduction to market devices. In *Market devices*, ed. M. Callon, Y. Millo and F. Muniesa, 1–12. Oxford: Blackwell.
- Nóvoa, A., and T. Yariv-Mashal. 2003. Comparative research in education: A mode of governance or a historical journey? *Comparative Education* 39, no. 4: 423–38.
- Oakley, A. 2002. Social science and evidence-based everything: The case of education. *Educational Review* 54, no. 3: 277–86.
- OECD. n.d. PISA – The OECD Programme for International Student Assessment. OECD. <http://www.pisa.oecd.org/dataoecd/51/27/37474503.pdf>.
- OECD. 2008. *Measuring improvements in learning outcomes: Best practices to assess the value-added of schools*. Paris: OECD Publications.
- Ozga, J. 2008. Governing knowledge: Research steering and research quality. *European Educational Research Journal* 7, no. 3: 261–72.
- Ozga, J. 2009. Governing education through data in England: From regulation to self-regulation. *Journal of Education Policy* 24, no. 2: 149–62.
- Ozga, J., and B. Lingard. 2007. Globalisation, education policy and politics. In *The Routledge Falmer reader in education policy and politics*, ed. B. Lingard and J. Ozga, 65–82. London: Routledge.
- Prickett, S. 2002. Conclusion: Metaphors, managerial ethics and unintended consequences. In *Education! Education! Education!* ed. S. Prickett and P. Erskine-Hill, 181–204. Thoverton, UK: Imprint Academic.
- Rizvi, F., and B. Lingard. 2009. The OECD and global shifts in education policy. In *International handbook of comparative education*, ed. R. Cowen and A.M. Kazamias, 437–53. Dordrecht, The Netherlands: Springer.
- Rose, N. 1999. *Powers of freedom: Reframing political thought*. Cambridge: Cambridge University Press.
- Rose, N., and P. Miller. 1992. Power beyond the state: Problematics of government. *British Journal of Sociology* 43, no. 2: 173–205.

- Scheerens, J., C. Glas, and S.M. Thomas. 2003. *Educational evaluation, assessment, and monitoring: A systemic approach*. Lisse, The Netherlands; Abingdon, UK; Exton, PA and Tokyo: Swets and Zeitlinger.
- Simons, M. 2007. 'To be informed': Understanding the role of feedback information for Flemish/European policy. *Journal of Education Policy* 22, no. 5: 531–48.
- Slavin, R.E. 2008a. Perspectives on evidence-based research in education: What works? Issues in synthesizing educational program evaluations. *Educational Researcher* 37, no. 1: 5–14.
- Slavin, R.E. 2008b. Evidence-based reform in education: What will it take? *European Educational Research Journal* 7, no. 1: 124–38.
- Smith, A.F.M. 1996. Mad cows and ecstasy: Chance and choice in an evidence-based society. *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 159, no. 3: 367–83.
- Smith, A.F.M. 2001. Public policy issues as a route to statistical awareness. *International Statistical Review* 69, no. 1: 17–20.
- Stengers, I. 2005. The cosmopolitical proposal. In *Making things public: Atmospheres of democracy*, ed. B. Latour and P. Weibel, 994–1003. Cambridge, MA: MIT Press.
- Strand, S. 1998. A 'value added' analysis of the 1996 primary school performance tables. *Educational Research* 40, no. 2: 123–37.
- Teddlie, C., and D. Reynolds, Eds. 2000. *The international handbook of school effectiveness research*. London and New York: Falmer Press.
- Van de Perre, L., M. Simons, and G. Kelchtermans. 2009. *Struggling with evidence: Knowledge and its paradoxal role in educational policy-making*. Paper presented at the ECER-conference, September 28–30, in Vienna, Austria.
- Venturini, T. 2010. Diving in magma: How to explore controversies with actor-network theory. *Public Understanding of Science* 19, no. 3: 258–73.