Curriculum Standards: Performance and Competence

Scott, David
Institute of Education, University of London
Email: d.scott@ioe.ac.uk

Abstract
The development and use of standards has become integral to educational practices in Europe, and makes possible certain forms of teaching and learning and excludes others. It works through fixing a particular representation of knowledge in those texts, tools, technologies, protocols, ethical desiderata and norms which constitute the educational practice, and it is increasingly being understood as global or at least extra-national. However, we should be aware of the inherent instability of standards and standards-delivery processes. In this article I examine current curricular forms in European countries by focusing on standards and the different models in use. I then contrast the dominant model (exemplified by the type of knowledge produced of an individual's or a group's capacities, i.e. nation, age-cohort or category, by international comparative systems of testing such as the Programme for International Student Assessment – PISA; and by quality assurance mechanisms), with a model that prioritises learning over performance.

Keywords: Curriculum, Curriculum Standards, Performativity, Assessment/Evaluation.

Introduction
Governments in Europe at the end of the twentieth century and in the early part of the twenty-first century, with a few notable exceptions, have reached a settlement about the nature of the school curriculum. This consensus now operates at all levels of the education system, and can be expressed in terms of a number of propositions: traditional knowledge forms and strong insulations between them need to be preserved; each of these knowledge forms can be expressed in terms of lower and higher level domains and the latter have to be taught before the former and sequenced correctly; certain groups of children are better able to access the curriculum than other children, and thus a differentiated curriculum is necessary to meet the needs of all school learners; the teacher's role is to impart this body of knowledge in the most efficient and effective way, and thus their brief cannot concern itself with the ends to which education is directed, but only the means for its efficient delivery; and the school's role is to deliver a public service that meets the targets set for it by governments.

In this article I examine current curricular forms in European countries by focusing on standards and the different models in use. Standards have become integral to educational practices in Europe, and make possible certain forms of teaching and learning and exclude others. They work through fixing a particular representation of knowledge in those texts, tools, technologies, protocols, ethical desiderata and norms which constitute the educational practice, and are increasingly being understood as global, or at least extra-national, phenomena.
However, we should be aware of the inherent instability of standards and standards-delivery processes. I then contrast the dominant model (exemplified by the type of knowledge produced of an individual's or a group's capacities (i.e. nation, age-cohort or category) by international comparative systems of testing such as the Programme for International Student Assessment (PISA) (OECD, 2001; 2006; 2009) and by quality-assurance mechanisms), with a model that prioritises learning over performance. Finally, I suggest that, with regards to standards and standards-based curricula, actor network theorists (for example, Fenwick & Edwards, 2010) offer a critique of these traditional forms, expressing their opposition in terms of claims about the emergent and unstable ontology of material and human objects, and the need to move away from prioritising intentionality and therefore human agency over other objects in the world. This undervaluing of agency and intentionality in the process of curriculum-making creates a particular problem for any theory of learning.

Standards and Standards-Based Curricula

There are four types of curriculum standards currently being used; all of which express standards in terms of statements of knowledge, skills and dispositions which it is expected students will have acquired at set stages during their time at school. The first model is where statements of curriculum standards are expressed at a high level of generality and abstraction and organised in a hierarchical order, so that there are progressively more complicated versions of each of the main ideas at each reporting stage (usually articulated in terms of more than the previous stage). The degree of generality and abstraction of the curriculum standards means that they can only act as a general guide to stakeholders (i.e. teachers, head teachers, examination constructors, policy-makers, parents and the like); and this therefore restricts the types of progression which can operate in the programmes of study. The end-result of this is that they are rarely used, and serve as adornments, rather than practical useable technologies.

A second model comprises a set of curriculum standards written as statements of expected achievements and at a level of concreteness which can be easily and reliably converted into useable products. These aim to capture the essence of the learning outcomes or objectives (sometimes written as competencies) of the educational process (and of course of the written curriculum, which is an attempt to capture this framing of the process). There is an element of reductionism in this model, as there is in any such model; any set of standards is organised in a more or less reductionist way. The element of progression between the different key stages is understood as having a number of dimensions and not just extension; in other words, there are different forms of relations between the different levels, such as: prior condition, maturation, extension, intensification, complexity, abstraction and articulation.

A third model is where the standards are written so that there is a high level of generality of the statements of expected achievements, with at the same time a more flexible conception of progression, understood in terms of a multiplicity of progression modes, a non-linear (in relation to the reporting stages) progression pathway, and the possibility that some skills, knowledge sets and dispositions may not feature at all at some of the key stages. The problem with this version of
the standards is that the very generality and abstractness of them may make their usefulness both more difficult and more unreliable. A fourth version is where assessment standards and curriculum standards are treated as equivalents. This has a number of problems: a reduction to what can be measured, a neglect of some standards which cannot be easily measured, and a possible distortion of the curriculum. In all four models, curriculum standards are understood as statements of expected achievements or level-descriptors defining expected achievements.

These four traditional models have a number of common characteristics: a conflating of curriculum and assessment/evaluation standards, a circumventing of the need for a pedagogic translation of all types of knowledge at the sites of learning, and a blindness with regards to washback effects, especially if we are dealing here with high status curricula and formal testing regimes. In opposition to these traditional models is a model which separates out curriculum standards from pedagogic standards, and even more importantly from assessment standards, requiring processes of translation to connect the three. In this model, curriculum standards are not the same as assessment standards. It is therefore important that the standard is not compromised in any way by whether it can or cannot be used as a testable construct or teaching approach. An assessment standard specifies those knowledge-sets, skills or dispositions (at different levels of difficulty), which a student is required to have, and which are expressed in such a way that they can be tested in a controlled environment, such as an examination. The principal problem with assessment standards is that testing a person’s knowledge, skills and aptitudes is likely to have washback effects on the original knowledge or skill set. Instead of the assessment acting as a descriptive device, it also acts in a variety of ways to transform the curriculum it is seeking to measure. Washback effects work on a range of objects and in different ways. So, for example, there are washback effects on the curriculum, on teaching and learning, on the capacity of the individual and more fundamentally on the structures of knowledge, though these four mechanisms are frequently conflated in the minds of educational stakeholders. Micro washback effects work directly on the person, whereas macro washback effects work directly on institutions and systems, which then subsequently have an impact on individuals within those institutions and systems. Finally, a student may have to reframe their knowledge or skill set to fit the test, and therefore the assessment of their mastery of this knowledge or skill is not a determination of their competence, but a determination of whether they have successfully understood how to rework their capacity to fit the demands of the examination technology. As a result teaching to the test occurs and the curriculum is narrowed to accommodate those standards which can be more easily assessed.

Pedagogic standards are instructional methods and range from didactic to imitative to reflective and meta-reflective approaches. In specifying a pedagogic approach or standard one needs to: specify the circumstances in which it can be used in the specific learning environment; specify the resources and technologies needed to allow that learning to take place; specify the type of relationship between teacher and student, and student and student, and student and parent, to effect that learning; specify a theory of learning – how can that construct (i.e. knowledge set, skill or disposition/inclination) be assimilated; and develop a
theory of transfer held by the teacher; that is, how can the learning which has taken place in a particular set of circumstances (i.e. a classroom, with a set of learners, in a particular way, with a particular theory of learning underpinning it, and so forth) transfer to other environments in other places and times. And again, it is now clear why curriculum standards need to be separated from pedagogic standards. First, they have different functions. Second, there needs to be a bridging mechanism between pedagogic and assessment standards for the reasons given above. Third, what is learnt, expressed in terms of knowledge, skills and dispositions, and is therefore applicable in a range of settings away from the classroom (i.e. in the future and in different environments than the classroom), is not the same as that which was originally learnt in school. This is where competencies (which are always future-orientated) come into the picture. However, to operationalise them, there is a need for a pedagogic theory or mechanism.

Another example of the need in this model for a pedagogic theory relates to the issue of summative and formative assessment. Curriculum standards can be used in a number of different ways, with different consequences. They can be used to determine whether and in what way the individual is meeting them, as well as providing information about how the individual can perform better in the future. (This requires a curriculum standard to be translated into a pedagogic standard.) Or they can be used to summarise levels of achievement at group, school or national levels. (This requires a curriculum standard to be translated into an assessment standard.) In summary, they can be used summatively or formatively. If these two functions are combined, then the curriculum is liable to be treated partially.

Programme for International Student Assessment (PISA)

The Programme for International Student Assessment (PISA) is both an example of knowledge production (in relation to standards) and a means for delivering those standards. In essence it is a standardisation process. PISA test constructors have chosen to measure competencies rather than knowledge on the grounds that knowledge is specific to particular countries, whereas competencies are universal. There are two problems with this. First, those national and local features of knowledge domains apply in equal measure to skills, competencies and dispositions. Second, there is a longer and more complex inferential chain involved in the measurement of competencies than there is in the measurement of knowledge acquisition, and there is therefore a greater likelihood of construct-irrelevance variance occurring.

PISA has attempted the difficult task of constructing curriculum-free tests; the most notorious example being the 11+ examination in the UK. The reason for this is that making comparisons between the test performances of students from different countries, with different curricula and with different teaching methods and approaches, requires the selection of test items that do not reflect national curricula or national pedagogic methods. So these international comparative tests are not a measure of their curriculum, nor what they have been taught, nor are they a measure of what students have learnt in any formal sense. This means that the content of the test items and their presentation are likely to favour some countries at the expense of others.
Cultural differences can be expressed in a number of different forms, such as, ascribing different values, and different strengths of values, to cultural items, or determining the nature, quality, probative force, relevance-value and extent of evidence, or focusing on practices which may be more familiar to people in some countries and less so in others. However, more importantly, cultural differences in relation to the selection of test items refer to the expression of the problem to be solved. If, for example, differences between national idioms, different national ways of thinking embedded in language forms, and different normic values woven into the fabric of national discourses are ignored, then the presentation of the actual test items as well as the range of possible answers that can be given may favour students from one nation at the expense of students from another.

This is a problem of fair comparison. And in order to make such a comparison, it may not just be a question of translating the words which are being used, that is, substituting one set (words, sentences, language structures) for another, but transposing the example and the problem, so that it now better reflects its new epistemic base. Underpinning the notion of an international test is the idea of a universal, i.e. culture-free, form of knowledge, which can be adapted so that superficial differences between nations are eliminated. However, it is never enough to say that a test simply tests the capacities and knowledge constructs of a group (in this case a trans-national group) of students. What a trans-national test does is make a number of reductionist assumptions about the knowledge bases being tested which result in imperfect representations of all the national knowledge bases under consideration.

PISA results are expressed as comparative national tables rather than scores achieved by participants. The focus is on position rather than score, even though significant improvements made by one nation between two time points may be masked by improvements made by other nations. If you add to this the idea that there is some uncertainty or unreliability about the scores (i.e. marker error, poor performance by testees, cultural bias effects, epistemic differences, inability to transform internal knowledge into performative knowledge, etc.), it is hard to believe that such league tables can and do provide a nation with very much useful information. However, what we have here is a display mechanism (located initially at the transitive level, but also penetrating and thus taking on a capacity to operate at the intransitive level). This display mechanism clearly has scientistic aspirations (cf. Habermas, 1987), adding further to the need to introduce critical and evaluative elements into any accounts made, whether they refer to individuals, groups within nations, or nations themselves.

PISA is a performative device, in so far as its intention is not just to describe the capacities of children but to promote and thus contribute to national policy-making. Certain forms of performative knowledge become the norm. The instrument for measuring knowledge and skill levels of children becomes an instrument for determining what those knowledge levels and skills should be, and how they should be learnt. In effect what the standardisation mechanism comprises here is an attempt to create the conditions for comparability and thus effectively the capacity to differentiate in a population of learners (i.e nations, groups within nations, and individuals). The mechanism that underpins this series of actions is an example of synchronic emergent powers materialism (cf. Bhaskar et al., 2010), and as a result, it operates as a standardising device in relation to
these matters (i.e. it creates a norm) and should not be understood as a device for making fair, reasonable and accurate judgements about the capacities of cohorts of students in different countries. There is a final point to be made, and this is that a nation’s place in these league tables becomes part of the folkloric account a nation gives of and to itself. Since this account is an important part of a nation’s identity, then success in an international test such as PISA becomes even more important. However, even here we should be aware of the inherent instability of these representations and mechanisms that we are suggesting operate within the trans-national educational practice.

**Bureaucratic Knowledge and Accountability Mechanisms**

Another example of standards development and deployment is through the production of bureaucratic knowledge, and in particular, the creation of forms of accountability within educational systems. Giving an account of a series of activities is not a neutral activity, but changes the nature of that activity, and acts to transform our understandings of it and thus our response to it. An example of this is those mechanisms set up to monitor teaching and learning processes in the United Kingdom. There is a disjuncture between the actual process of learning and those technologies which are both intended to allow that learning to take place in a more efficient manner and monitor the effectiveness of that learning. The disjuncture occurs because these technologies contribute little to the process of learning; in effect, they are different activities with a different focus, though they purport to be about the same issue. Quality assurance mechanisms have as their purpose an intention or desire to change what is happening in the practice, but this is because they act in an ideal sense so the teacher conforms imitatively, or is compelled to conform or comply because of a fear of sanctions, or because those sanctions have been applied. What frequently results is a simulation where the teacher conforms on the surface to the demands of the quality assurance process, but in fact operates through a different set of logics. Whether they do this successfully is a different matter because they have to be highly skilled in playing both games simultaneously; in effect operating discursively along parallel tracks and making sure that the one doesn’t affect the other. Their sense of direction however, is always primarily towards putting in place the optimal conditions for learning of their students. Though the purpose of bureaucratisation is to act as a form of labour control, this term fails to give expression to the full import of the process, because the colonising process achieves its purpose through changing the epistemology of the setting. This entails a displacement of content through operating a standardised bureaucratic form of knowledge.

This is an example of a quality assurance mechanism in operation, though this mechanism has a variety of forms, which can be understood as positions on a number of scales: the degree to which they engender a low or high level of trust within the system; the degree of punitive strength they can muster; their capacity to influence the activities under scrutiny, for example, whether they can or cannot initiate washback effects; their capacity to influence the epistemological character of the setting; the degree of affordance they give to participants in the setting; and their underpinning ideological framework regarding human nature and possible forms of human interaction.
Most quality assurance mechanisms ask for explicit rather than tacit accounts of practice, and this is not just a question of showing or demonstrating, but of the practitioner stating them in a formal codified way (and in a particular form which means that they have particular consequences). Here there is direct engagement with the constitutive practices. However, the nature of the practice may be distorted by the desire to make it explicit; in other words, there may be a problem of reductionism, especially if the language used is quantitative.

An evaluator or standards-setter has to, in the first instance, make an epistemic judgement about the boundaries of the programme of teaching and learning activities they are investigating. Standardisation therefore has an active element to it. Since a programme by definition includes activities-over-time, then that judgement needs to take account of changes to the programme which are caused by the actions of internal or external participants. As a consequence, this initial epistemic judgement is in fact a series of judgements (i.e. $J_a$, $J_b$, $J_c$ .... $J_n$), and, in addition, the judgements themselves as parts of a series may not be in line with the boundaries of the object being scrutinised, i.e. the programme of teaching and learning. The next step is making a judgement about the programme; and this requires two types of judgement to be made: the first is about the type of evidence required to make a judgement and the second is about the type of inferential relationship that is present between evidence set and conclusion. And finally, the evaluator or standards-setter makes a value-laden judgement about the programme’s activities, which refers firstly to their own set of values and secondly to the type of judgement that they choose to make.

All judgements about educational matters are inferential judgements about evidence and the conclusions that the investigator wishes to draw. In making a judgement about a system (whether it meets the pre-specified standard), an institution (whether those pre-designated characteristics of an ideal form of such an institution are in place) or the performance of a person (whether the appropriate curriculum standards have been met), evidence and its analysis are central. There are two types of evidence: primary data, which is not and cannot be a-theoretic, and comes in the form of testimony or direct observations of worldly events or happenings; and a codified chain of reasoning which involves the collection and analysis of primary data and the positioning of those data in an inferential sequence to allow a conclusion or judgement to be made (as to whether and to what extent an hypothesis about the organisation or person is reliable and valid). Evidence can be more or less authentic, reliable and accurate, and more importantly, more or less salient, where this is defined as a chain of reasoning involving evidence and inference leading to a conclusion about a set of activities and involving judgements at every level. So, a piece of evidence may have a weak indirect relationship to the chain of reasoning, or a strong direct relationship to the chain of reasoning, because it refers to the chain itself and not to evidential elements of it.

Furthermore, salience as a criterion for determining the suitability of a piece of evidence for supporting a judgement is practice-specific. This refers to the kinds of information which serve as supporting facts in making a claim, and these, it is suggested, are practice-dependent: what is a relevant fact is determined within a practice. Therefore evidence may not be relevant because it does not fit with the evidence base which that claim is embedded within and
which gives it some measure of credibility. And further to this, each and every

evidence-set also has within it a threshold for determining the required probative
force of any claim that is made.

Evidence in relation to a judgement about a system, institution or person
therefore may be illegitimate for a number of reason: domain incommensurability;
non-conformity to the implicit and explicit rules of the domain; the degree of
probative force to achieve credibility within the domain; whether it fits with the
way the domain is formed; the degree and type of fallibility accepted in the
domain; and the degree to which the evidence set provides a complete or
incomplete account of the activities being investigated. The content and form that
evidence might take differs between domains. And this in turn means that
judgments which relate to other domains are illegitimate when applied to
particular domain-specific sets of evidence and inference.

There are a number of ways by which such judgements could be made. The
first is deontological, where the judgement is made in terms of a set of
absolutely right actions or a set of universal precepts. A second way is
consequentialism. This suggests that a judgement is made in relation to the
consequences of the actions of participants in the programme, and not in terms
of intention, circumstance or process. A third way by which such judgements can
be made is by examining the intentions of the programme or person, and then
comparing what has actually happened with what was intended to happen. There
are a number of problems with this. Intentions are always future orientated, and
fundamentally they reflect what key participants think can be achieved in relation
to what currently exists and how what currently exists may change in the future,
i.e. they are predictive. However, they may be wrong, misguided, badly
formulated, incorrectly predictive, etc. The question has to be asked: who is
responsible for the programme?

In making these judgements the evaluator of the programme or standards-
setter (whether internal or external) is making a judgement about the amount and
type of moral responsibility which can be attributed to those social agents who
are central to the activities of the programme. This raises a number of questions
about moral responsibility: Does the person (or persons) qualify as a moral agent
(or moral agents)? Do they possess the general capacity to perform as a moral
agent, where this refers to an ability to evaluate their reasons for doing this rather
than that? Are the conditions in place in the setting which is being evaluated that
allow the agent to perform in a way that conforms to their sense of moral
accountability, i.e. have they performed it freely and were they allowed to
exercise their moral culpability? And finally, have they taken sufficient account of
the conditional nature of any decision-making they might want to engage in?
(This conditionality has four forms: social actors are relatively unaware of some of
the conditions for their actions; that is, every action has a set of conditions
underpinning it, for example, a speech act requires a language, vocabulary and
grammar; they are unlikely to be able to predict all the consequences of their
actions, so there are going to be unintended consequences; social actors may
not be aware of much of their own knowledge and expertise, in other words,
much of their knowledge is tacit, and thus they cannot, except with the greatest of
difficulty, surface it in their accounts of their lives; and equally they may be
motivated by unconscious forces and impulses which they find great difficulty in
articulating. A distinction can be drawn between attribution and responsibility as accountability (Aristotle, 1925), and this distinction rests on the difference between ascribing moral responsibility to a person or organisation because they or the organisation is formally responsible for their or its activities and only making someone or some organisation responsible if they were in a position to do something about it and thus effectively make a difference. This last involves a judgement about what is reasonable in attributing praise or blame to a person or organisation in the actual circumstances in which those activities were performed and about which that judgement is being made.

Conclusions

The development and use of standards has become integral to educational practices in Europe, and make possible certain forms of teaching and learning and exclude others. They work through fixing a particular representation of knowledge in those texts, tools, technologies, protocols, ethical desiderata and norms which constitute the educational practice, and are increasingly being understood as global or at least extra-national. I have illustrated these processes of standardization by describing two delivery mechanisms: international comparative testing systems and quality assurance processes. However, we should be aware of the inherent instability of standards and standards-delivery processes. With regards to standards and standards-based curricula, actor network theorists (for example, Fenwick & Edwards, 2010) have focused on the instability of these traditional forms, expressing their opposition to our conventional understandings of these terms by pointing to the emergent and unstable ontology of material and human objects, and the need to move away from prioritising intentionality and therefore human agency over other objects in the world. This undervaluing of agency and intentionality in the process of curriculum-making creates a particular problem for any theory of learning and standards development.

Determinism would imply in its strongest form that our thoughts, feelings and subsequent behaviours do not deviate from the impulsions laid down in our genetic make-up or in customised knowledge within our bodies or in the social arrangements (i.e. embodied, discursive, agential, institutional and systemic) that constitute our lives. If we want to build in a notion of agency, then we have to believe that our cognitive and volitional capacities can operate without recourse to and outside of those causal impulses that come from these determining impulses. Furthermore, if we hold to a belief that our cognitive and volitional capacities are inextricably tied to our genetically-determined, embodied or socially-determined impulses, then it follows that our capacity to determine whether or not we are being deceived, i.e. our capacity to tell the truth or not about our fundamental belief in determinism, is thoroughly compromised. Agency therefore involves a set of activities which are not caused or influenced by those causal impulses that emanate from our genetic, embodied or social impulses; that is they do not involve a negation of them or a reaction against them. What this means is that agency is central to learning and standard-setting, and due regard should be paid to such a notion in any curriculum deliberation.

Basil Bernstein (2000) identified two models of curriculum and called these performance and competence, with the former now the dominant model in
Europe. The performance model has its origins in the behavioural objectives movement, and though contested by curriculum theorists, retains its status as the dominant model. It is a model which clearly emphasises marked subject boundaries, traditional forms of knowledge, explicit realisation and recognition rules for pedagogic practice, and the designation and establishment of strong boundaries between different types of students. Such a model in the hands of policy-makers becomes both normative and teleological. Bernstein compares this with a competence model, and in relation to the latter, he suggests that acquirers have some control over the selection, pacing and sequencing of their curriculum. Performance modes are now seen as the norm, whereas competence modes are seen as interruptions or resistances to the normal flow of events and this has in part been achieved through processes of standards-setting and standards-development set out above.

References

Received: 23 October 2013
Accepted: 10 January 2014