

Primary education curricula across the world: qualitative and quantitative methodology in international comparison

Dominic Wyse & Jake Anders (UCL Institute of Education)

Accepted Pre-Print Version

Provision for children from age five to age eleven is a vital phase in formal education settings worldwide. During this *primary education* phase (also called *elementary education* in some countries) children should make the journey from acquiring basic understanding to developing the bases for ways of thinking that will be used throughout their lives. Primary education is also an essential preparation for secondary education. The international importance of primary education is recognised by the United Nations Educational, Scientific and Cultural Organization (UNESCO) as part of their “17 goals to transform our world: Goal 4: Ensure inclusive and quality education for all and promote lifelong learning.” (UNESCO, 2017, online) However, in a report in 2017 it was found that “more than 387 million children of primary school age (about 6 to 11 years old) and 230 million adolescents of lower secondary school age (about 12 to 14 years old)” were not achieving minimum proficiency levels in reading and mathematics (UNESCO, 2017, p.2).

The idea of proficiency in any subject area is strongly related to the curriculum, in the activities that children encounter in their classrooms, in the organisation of learning throughout schools, and hence in the specification of countries’ national curricula. The curriculum that children experience in formal education settings is one of the fundamental elements that determines the nature of their learning. Reading and mathematics are important aspects of any primary curriculum, and are subject to the most common comparative tests, but the nature of the whole curriculum, its aims and areas/school subjects has to be taken into account when seeking to understand the effectiveness of national curricula.

This chapter begins by exploring definitions of curriculum and linking this with the place of knowledge in relation to curriculum studies. The chapter reveals that although knowledge is a prime focus of the curriculum studies field it is far from clear how this might guide the most appropriate content for modern curricula in the range of cultural contexts that are represented by different countries. Having established a theoretical frame the main part of the chapter features an analysis of the traditions and methodology of qualitative comparison of countries’ curricula compared with quantitative measures of different curricula. One of the most impressive qualitative comparative studies of recent years is used as the basis for a detailed comparison of findings and methods for international comparison. This section of the chapter leads to reflections on the representation of oral language development as part of national curricula and classroom practice, as identified in this seminal study. Qualitative research methodology to explore primary curricula, which underpins this study, is then compared with quantitative methodology represented by international comparative work. This section of the chapter

features some examples of curriculum issues such as the teaching of reading that are represented in the questions that are asked in surveys that are part of the methodology.

The examples chosen to illustrate curriculum comparison and its methodology are all representations of the complexities of how knowledge can be selected and represented in curricula. For example the presence, or lack of presence, of oral language in curricula, is compared with the ways in which the teaching of reading (which draws on oral language competence) is represented. Strengths and limitations of both methodologies are considered with the aim to suggest stronger designs in future. The chapter concludes with some suggestions for future research possibilities for the comparison of countries' curricula.

Curriculum and Knowledge

Definitions of curriculum differ according to the disciplinary context, and by the scope and scale of the conception of curriculum. For example, from a sociological perspective curriculum has been defined as “the principle by which units of time and their contents are brought into special relationship with each other” (Bernstein, 1971, p. 48). From an educational perspective curriculum can be, “What is intended to be taught and learned overall (the planned curriculum); what is taught (the curriculum as enacted); what is learned (the curriculum as experienced)” (Alexander, 2010, p. 250). The sociological definition, principles of the relationship between units of time, is succinct but does not account for the place of pedagogy, including the teacher-pupil interaction that is an essential part of curriculum. The educational definition with its distinction made between curricula as intended, enacted and experienced is helpful although enactment and experience are so closely related this makes the utility of this definition for analysis purposes difficult. As a result, the focus in this chapter is on curriculum as “planned human activity intended to achieve learning in formal educational settings” (Wyse, Hayward and Pandya, 2016).

Curriculum studies, as a field of the discipline of education, has established the importance of questions about *knowledge* in curricula. This importance is underlined by the idea that the specification of knowledge, realised through classroom activity, is a curriculum's main function. As early as the 19th century Herbert Spencer asked, what knowledge is of most worth? Putting aside the dubious introductory comparisons with indigenous people's customs at the beginning of Spencer's book, it is remarkable, in view of the time of publication, to see Spencer's explicit attention to rationales for knowledge in the curriculum:

If there needs any further evidence of the rude, undeveloped character of our education, we have it in the fact that the comparative worths of different kinds of knowledge have been as yet scarcely even discussed – much less discussed in a methodic way with definite results. Not only is it that no standard of relative values has yet been agreed upon ; [sic

– spacing] but the existence of any such standard has not yet been conceived in any clear manner. And not only is it that the existence of any such standard has not been clearly conceived ; but the need for it seems to have been scarcely even felt. [semi colons and their spacing as in original text] (Spencer, ND, published 1860, p.11).

However, the contribution of curriculum studies has in the last decade or so been characterised as in ‘crisis’. The reasons for the crisis have at least three possible roots: 1. A perception of domination by “empiricists” (who are typified as basing their claims for particular curricula on international testing and its analyses) and “post-conceptualists” (with an emphasis on personalising concepts derived from post structuralism and German existentialism) at the expense of “traditionalist work based on close relations with teachers’ work and/or curriculum development.” (Hopman, 2010, online). 2. theoretical neglect of access to knowledge as part of the curriculum. And more particularly,

the reluctance of curriculum theory, at least since Hirst and Peters (1970), to address epistemological issues concerning questions of the truth, and reliability of different forms of knowledge and how such issues have both philosophical and sociological dimension (Young, 2013, p. 103)

3. The crisis has also more recently been located in the nature of political control of the curriculum, and in particular, politicians’ perceptions of risk associated with international league tables of pupil testing which have led to greater intervention in national curricula. The pressures caused by high stakes testing have been seen as a recent manifestation of *performativity* (Wyse, Hayward & Pandya, 2016).

In its original conception the French philosopher Lyotard saw performativity as revolving around the perceived need for two sets of “skills” (in relation to university education rather than schools and early years settings). One set of skills, he argued, was needed for selling on the world market through prioritisation of ‘particular subjects to support growth in demand for high and middle management executives; the other set of skills was those needed for maintenance of internal cohesion, a strategy that pushes out other aims of university education such as those built on emancipatory narratives (Lyotard, 1984, p. 50). This performativity brought “inevitable disorders ... in the curriculum, student supervision and testing, and pedagogy, not to mention its sociopolitical repercussions” (*op cit.*). Lyotard concluded that in a performativity culture educational institutions are therefore subordinated to the ruling powers.

Although Lyotard’s analysis was applied to universities, we can see a possible link with schooling. The significance of selling on a world market can be seen in globalisation trends (typified by global networks with growing influence on education policy, Ball, 2012); in the finance and influence of international comparative testing; and the associated influence of international testing on subjects in national curricula such as literacy and mathematics. Greater

control over curricula, and increasingly pedagogy, can be used to maintain government ideology (Lyotard's 'internal cohesion') at the expense of student, teacher, and local authority control of education (the emancipatory narrative). And the disorders in curriculum, and socio-political repercussions, can be seen in the criticisms evident in academic analyses, particularly from the qualitative comparative tradition (for a very strongly worded example see Alexander's (2011) paper, "Evidence, rhetoric and collateral damage: the problematic pursuit of 'world class' standards").

The idea of ownership of the curriculum (Lyotard's emancipatory narrative) is in some ways in tension with current emphases on knowledge in curricula (Young, 2008), and hence a central problem for the curriculum studies field. Johan Muller, who has worked with Michael Young, suggested a possible way forward:

It seems that Wyse et al.'s view that knowledge is "both constructed and real" (2014: 5) was right after all. Quite how to establish the reality of "powerful knowledge" while acknowledging its social roots remains a challenge in 2014 as it was in Mannheim's day. What is undeniably underway is a sort of rapprochement, but it remains a work in progress. (Muller, 2016, p. 103)

The idea of knowledge being both constructed and real was a reference to the editorial of the special issue of the *BERA Curriculum Journal* (Wyse, et al. 2014, which highlighted Biesta's 2014 award-winning paper addressing the idea of transactions, and the idea of knowledge being both constructed and real). In addition to recognising the contribution made by Wyse et al., Muller's suggestion about rapprochement was built on a series of other broad points. One part of his argument was a reminder of the significance of the work of Basil Bernstein carried out while at London's Institute of Education (since 2015 part of University College London UCL). Bernstein was critical of some theories of cultural reproduction that had emerged from France. He argued that their conception of education as a "carrier of power relations" such as "class, patriarchy, race" (Bernstein, 1996, p. 4) resulted in a lack of attention to "internal analysis of the structure of the [pedagogic] discourse itself, and it is the structure of the discourse, the logic of this discourse, which provides the means whereby external power relations can be carried by it." (*op cit.*)¹ However, Bernstein may not have paid due attention to Lyotard's depiction of the legitimization of education through performativity which included the "transmission" of an established body of knowledge. This led Lyotard to a series of pragmatic questions: "Who transmits learning? What is transmitted? To whom? Through what medium? In what form? With what effect?" (p. 48). It is true that Lyotard's analysis portrayed education as a carrier of power relations, such as forces of performativity, yet Lyotard did appear to address some of the discourse features as well.

¹ It is important to note here that Bernstein's interest in the pedagogic discourse of curriculum was distinct from his preoccupation with the language of pupils, in particular his dubious articulation of *restricted* and *elaborated* code.

For curriculum theorists, Bernstein's distinction between *classification* as "the degree of boundary maintenance between contents" (1971, p. 46) and *framing* as "the degree of control teacher and pupil possess over the selection, organization and pacing of the knowledge transmitted and received in the pedagogical relationship" (*op. cit.*) has been significant.

Where framing is strong, the transmitter has explicit control over selection, sequence, pacing, criteria and the social base. Where framing is weak, the acquirer has more apparent control (I want to stress apparent) over the communication and its social base. (Bernstein, 1996, p. 12).

For example, a subjects-based curriculum (which Bernstein called a "collection code" curriculum) has strong classification, whereas a theme-based curriculum (an "integrated code" curriculum) has weak classification. Bernstein identified "some reasons for a movement towards the institutionalising of integrated codes *of the weak classification and framing (teacher and taught above the level of the primary school)*" (Bernstein, 1971, p. 66 – italics in original)". In national curriculum terms, weak classification might be realised as a thematic curriculum (such as the International Baccalaureate IB, or an aims-based curriculum - Reiss and White, 2013), and weak framing might entail strong control at teacher-pupil level rather than governmental level.

Bernstein's reasons for recommending moving towards an integrated code curriculum included recognition that: higher levels of thinking were increasingly differentiated, and that more flexibility was required in the labour force hence students needed empowerment to pursue their interests within this wide range; and that there was a need for more egalitarian education, not least to make sense of major societal problems related to power and control. Bernstein's thinking focused on ways to "declassify and so alter power structures and principles of control; in so doing to unfreeze the structuring of knowledge and to change the boundaries of consciousness." (Bernstein, 1971, p. 67) He theorised that only a select few pupils/students are normally allowed access to "relaxed frames", in other words a state of empowerment for these pupils to "create endless new realities" as part of the understanding that knowledge is permeable and provisional.

Bernstein's proposal seems to differ in important ways from Young's proposals for implementation of a subject-based curriculum, as can be seen in some of Young's work:

5.2. The relationship between a national curriculum and the individual curricula of schools

A National Curriculum should limit itself to the key concepts of the core subjects and be designed in close collaboration with the subject specialists. This limit on National Curricula guarantees autonomy to individual schools and specialist subject teachers, and takes account of schools with different cultural and other resources, different histories

and in different contexts (for example, schools in cities and rural areas). At the same time, it ensures a common knowledge base for all students when some may move from school to school. (Young, 2013, p.110)

Young's emphasis on "core subjects" (in schools) and "subject specialists" (presumably in universities) represents strong classification, and although governmental limits on the content specification of national curricula might enable "autonomy" for individual schools, and might appear to represent weak framing, the pressure to ensure the place of the core-subjects (in systems where national testing is 'high stakes') is likely to be at the expense of both pupil and teacher autonomy and at the expense of subjects beyond the core (Wyse and Torrance, 2009).

International Comparison and Primary Education Curricula

The beginning of international comparison of countries' education systems and curricula is often attributed to Marc-Antoine Jullien (1775-1848). In his "Plan and preliminary view for a work on comparative education", originally published in French, Jullien argued for the importance of comparison of the education systems in Europe in order to improve them and contribute to societal avoidance of the horrors of future wars. When viewed in relation to current debates about international comparison of education systems, Jullien's suggestions were visionary:

In order for educational science to keep up, spread, and perfect itself, it, like other sciences, requires many nations at the same time to interest themselves in it and practice it together. Competition becomes useful to those very ones who would at first think to see in it an obstacle to their interests. A wise and well-informed politician discovers in the development and prosperity of other nations a means of prosperity for his own country. (Fraser, 1964, p. 37)

Jullien's emphasis on "science", "competition", "prosperity", and the link with politics, has strong resonances with contemporary debates about international comparison of education systems.

Jullien's plan included specifications for surveying education systems in different countries. See for example some of Jullien's categories of questions:

FIRST SERIES [of questions]

A. PRIMARY AND COMMON EDUCATION

1. Primary schools or elementary and common
2. Directors
3. Students
4. Physical education and gymnastics
5. Moral and religious education

6. Intellectual education and knowledge
7. Domestic and common education, as it is related to public education
8. Primary and common education, as it is related to secondary education or to the second stage, or with the intentions of children.
9. General considerations and various questions (*op cit.* p. 50)

In category 6. *Intellectual Education* Jullien's suggested questions include these:

91. How does one conduct from the cradle, the first education of senses and organs? With what objects is care taken to surround children, to exercise them to see, touch, hear, taste, feel? What are the first exercises of observation and language?

92. At what age are children usually taught to read, write, count, and what method is considered the easiest.

93. What are the aims of education which the children usually receive in primary school? (Does one limit oneself in the majority of these schools to reading, writing, arithmetic? Or does one also give a few elementary ideas of *grammar, singing, geometrical drawing, geometry, and land surveying, applied mechanics, geography and history of the country, anatomy of the human body, practical hygiene, natural history applied to the study of land products most useful to men?* All the elements of these sciences, as essentials to each individual in all conditions and circumstances of life, would seem to have to form a part of a complete system of primary and common education, perfectly appropriate to the true needs of man in our present state of civilization.) (*op cit.* p. 63)

The appreciation of the risk of a preoccupation with a narrow curriculum of the basics of reading, writing and arithmetic (the 3 Rs) predates modern debates, and Jullien also raised the importance of the holistic aims of primary education.

In more recent theoretical work, and in the qualitative tradition, various studies can be seen as arising out of Jullien's foundations, though not necessarily related to his appeal to science. For example, using curriculum as the unit of analysis Forestier and Adamson (2017) root their critique of PISA in Jullien's ideas in their review of comparative studies. Their critique notes the dangers of a narrow focus on literacy and mathematics; the need for a holistic and contextual approach to comparison, such as that taken by Alexander (2001); the need for an investigative orientation to comparison versus the evaluative and formative. However, although the paper is focused on curriculum it does not cite key work in the curriculum studies field, preferring to draw on a wider comparative tradition.

In another evaluation of the impact of PISA, Carvalho and Costa (2017) review six papers that summarised nations' responses to PISA results as follows: In France the focus was on what is legitimate knowledge in the curriculum for the functioning of education; in Hungary the PISA findings were used as a master narrative for policy; in Francophone Belgium the programme contributed to moves from regulated to deregulated state; and in Portugal it was used as a national evaluation tool, and to legitimate government policies. So, it can be seen that outcomes generated as a result of PISA are used to fulfil different purposes: to legitimize policy; manage the policy agenda; develop secondary research; and support development of domestic regulatory instruments.

Perhaps the most significant study of primary education in the qualitative comparative tradition is the *Five Cultures* project (Alexander, 2000). The ambition of this study is clear from the title of the work which sought to establish robust evidence about the culture and education systems of five nations. Indeed, it is made clear in the work that the context of the research was comparison of 'five continents'. The main publication resulting from the work is very extensive, and impressively multi-layered account, addressing several significant levels of education systems, from documentary analysis of national policy down to observations of teacher-pupil interactions and classroom practice.

The work is also explicit about the risks of generalisation from such studies. However, having exhaustively and memorably portrayed the five countries, the book singles out one particular finding to conclude definitively that education in England is deficient in comparison with some of the other countries:

Further, the prescribed English language curriculum for the primary stage - and indeed the secondary as well - makes far less of the development of 'spoken' language than do the equivalent statements in France and Russia ... This difference in emphasis is really quite striking, and it manifests itself in the characteristically episodic lesson structures with their relatively fast interactive and cognitive pace that I described earlier [in France and Russia]. The quality and power of children's spoken language gain immeasurably from this approach, as one would expect. Further, there is no evidence that the development of children's reading and writing are in any way disadvantaged as a result. On the contrary, the relationship and function of each seem to be better understood. (op cit. p. 565)

This conclusion is preceded by an outline of the place and nature of oral language, or *Speaking and Listening*, in the different versions of England's national curriculum. Alexander appropriately, in our view, identifies a deprioritisation of oral language in preference to written language over the different versions of England's national curriculum since 1988, a view warranted by documentary analysis of national curriculum texts. The argument is that with regard to oral language the emphasis in the programmes of study was reduced between the period from 1988 (when

England's first national curriculum was established) to the national curriculum of the year 2000. Although Alexander's original work could not cover this, the process of deprioritisation of oral language was significantly accelerated in England's national curriculum of 2014 (Wyse, Jones, Bradford and Wolpert, 2018). Evidence for Alexander's conclusion can be seen as linked with his coverage of previously published empirical studies carried out in England, for example to the work of Maurice Galton and team in relation to teacher-pupil interaction. But if similar empirical work has been carried out in the other four countries in the Five Cultures study this is not cited. Nor was any kind of systematic review of studies on oral language in different countries carried out, although this is understandable in view of the already large-scale nature of the Five Cultures project.

With regard to knowledge in the curriculum, language, particularly spoken language, is a good test of theories of curriculum, knowledge and comparative methodology because language is fundamental to all learning and as a result occupies a complex space in national curricula. It is known that oral language is fundamental to acquiring reading and writing. But mother tongue oral language is acquired naturally by nearly all children as a result of the innate characteristics of human beings, supported by the interaction of significant others in the child's life, such as parents/guardians (Goswami, 2008). There are some facets of oral language that can be seen as 'knowledge' in relation to the definition offered at the beginning of this chapter, for example the acquisition of vocabulary, the special use of the voice in drama, sensitivity to formality in social contexts, but oral language does not sit easily in relation to Young's idea of "a common knowledge base" nor even to Bernstein's idea of classification and framing of subjects. This is because oral language is predominantly the vehicle by which the curriculum is delivered much more than an area of subject content. As such there is an argument, contrary to Alexander's, to suggest that oral language might *not* need the extent of prescribed curriculum content as reading and writing, while at the same time recognising its fundamental link with literacy and the necessity for oral language to have an appropriate specification in curricula not least in relation to standard forms of language such as standard English (but for the complexities of the development, including historical development, of standard English see Wyse, 2017).

Alexander continues with his main conclusion, about oral language:

Close analysis of all the videotapes and transcripts from the Five Cultures project - which of course far exceed the few examples contained in Chapter 16 - force me unambiguously to the conclusion that in English primary classrooms, although much may be made of the importance of talk in learning, and a great deal of talking goes on, its function is seen as primarily social rather than cognitive, and as 'helpful' to learning rather than as fundamental to it. (Alexander, 2000, p. 566)

The conclusions from the Five Cultures study include a statement suggesting causation not just correlation: "The quality and power of children's spoken

language gain immeasurably from this approach, as one would expect.” (p.565). And the conclusions include a generalisation at the whole country level, “in English primary classrooms, although much may be made of the importance of talk in learning, and a great deal of talking goes on, its function is seen as primarily social rather than cognitive, and as 'helpful' to learning rather than as fundamental to it.” In order to provide evidence to support the first conclusion ‘quality and power’ would need to be defined then a robust measurement of oral language established. For the second conclusion, first and foremost the very broad and interpretable concepts of ‘social’ and ‘cognitive’ need to be clearly specified. Then, as the emphasis is on ‘perceptions’ of the function of oral language, a sufficient number of people reporting their perceptions would need to be established. What’s more, in the cases of both conclusions the methodological requirements to establish causal and generalisable conclusions would need to apply in each of the countries that are part of the study. And hence this raises questions about the methodology of such studies.

A key parameter for the Five Cultures study, and generally for quantitative and qualitative research comparing different countries, is sampling. The first sampling decision made was of course in relation to the five countries. Paradoxically, in view of Alexander’s criticisms of quantitative comparative work, the choice of countries is based firstly on numbers: five countries is better than two, it is argued, because it avoids a tendency to polarisation. Five is better than three because this avoids the “Goldilocks effect” (*op cit.* p. 44). Ultimately though the selection of countries was made because “the countries offer similarities, contrasts, and intriguing connections”. But surely any selection of countries, from two countries to 33 countries or more, would offer these connections? Perhaps the selection is most accurately described as ‘a convenience sample’, or ‘a non-probability sample’, a sampling choice that can be argued to be acceptable, but this is not stated explicitly nor the specific reasons why for example India not Pakistan (in relation to Britain’s colonial past), Russia not China (in relation to globalisation, colonial legacy and evolution). If there were particular contacts and/or networks, or other practical reasons, that facilitated the selection of countries it would have been helpful to know these.

The argument that comparison of education systems requires inclusion of observations of classroom practice is well made in *Culture and Pedagogy*, but the sampling of schools and classrooms in the Five Cultures study is not entirely clear in some respects. It appears that thirty schools divided between the five countries in the study were the basis for the comparison of lessons: “The Five Cultures data include material from fieldwork in thirty school buildings in five countries and from an additional sixty or so English schools [data from previous studies by Alexander]” (*op cit.* p. 177). The argument is made that by analysing schools, national policy, *and* histories of education in the different countries the problems of generalising from a small sample can be avoided. But in order to make a robust claim about *perceptions* of talk as primarily social, or primarily cognitive, in relation to a whole country many would argue that this would require a nationally representative sample of people to give their perceptions.

Although transcriptions of teacher-pupil interaction are a well-regarded means to assess one important aspect of classroom pedagogy (and too often not part of the methodology of studies comparing education systems) the numbers of teachers (and pupils) involved in the sampling is again important. The number of teaching sessions that were observed, annotated and recorded was as follows:

“France 20; India 19, Russia 33, United States 19, England 75 (60 from preceding projects together with 15 Five Cultures updates). This gave a total data set of 166 lessons. Of these, 36 (six to nine from each country) were selected for transcription and close scrutiny ... although any logistical generalisations below are based on the full range from each country.” (*op cit.* p. 276).

Apart from the lack of contextual background to the chosen lessons (such as teachers’ gender; experience; languages spoken; qualification level; or class sizes etc.) the sampling raises questions about the extent to which generalisation about emphasis on oral language in the different countries can be supported by the data. There is no information about how the selection was made for “close scrutiny”, nor about data analysis techniques used to arrive at the findings. For example, if “the full range” of videos of lessons was used how did the analysis account for transcribed lessons versus those that remained only in video form? Or, how was ‘fast pacing’ defined and measured? An alternative approach to the methodology could have included a more explicit account of the rationale for sampling. For example, in relation to the sampling of countries a different approach could have included more careful matching of countries by population size, main language spoken, and geography in order to sharpen the comparison of policies and classroom practice. In relation to the sampling of schools these might have been selected using a form of stratified random sampling in order to mitigate some threats to validity that are part of other means of selection, and to try and match some of the school contexts across different countries. And, as we argue at the end of this chapter, a more radical approach could have been to build on evidence from large-scale international comparative quantitative work with in-depth qualitative work, or the reverse.

Important findings from in-depth analysis of the contexts and cultures in different countries in comparison with other countries have been established. However, the challenge of generalisation remains, not least when seeking to establish an evidence base for the development of national curricula. Having examined a particularly notable qualitative study we now turn to quantitative work that compares countries and curricula.

Quantitative comparison of the five countries

As documented extensively elsewhere, including in this volume, internationally comparative quantitative analysis in education has grown rapidly over recent years, particularly because of the increasing availability of datasets designed explicitly for this purpose. The most prominent of these are

the OECD's Programme for International Student Assessment (PISA), Trends in International Mathematics and Science Studies (TIMSS) and Progress in International Reading Literacy Studies (PIRLS). Although these studies are primarily focussed on student achievement (albeit in different ways), and have their flaws, we believe they have the potential to contribute to understanding of differences between primary curricula across the world.

Although, as will be seen, direct comparison of the place of oral language in different curriculum is not possible using the quantitative data sets, we carry out some illustrative comparison using the countries included in the Five Cultures study. It is worth noting at the outset that one of the benefits of quantitative enquiry is to allow for analysis of a wider range of countries than is feasible through qualitative work. Nevertheless, decisions to compare a tightly defined set of countries available in internationally comparative data are common, for example on comparing economically developed Anglophone countries (e.g. Washbrook et al., 2012; Jerrim et al., 2016) because of the cultural elements, particularly language, that such countries share. As language or languages are such an important cultural element of any country or region, and such an important part of the curriculum at primary education level, the comparison of countries with the same dominant language can allow for more meaningful comparison. Such decisions will, of course, depend on the specific research question.

With the exception of India all but one of the countries featured in the Five Cultures studies have participated in PIRLS or TIMSS in recent years. This presents an opportunity to explore primary curricula in these countries that will complement analysis such as that discussed above, through use of quantitative data collected as part of these studies since 2003.

The most prominent element of the large-scale quantitative studies is the overall country rankings that are produced each time the results are published. As we demonstrated earlier these publications have increasingly attracted different kinds of responses from governments. With regard to the four Five Cultures studies included in PIRLS 2016 their rank order is shown in Table 1.

Table x.1: Extract from PIRLS 2016 Reading Achievement Distribution (IEA, 2016a)

				5th Percen tile	25th Perc entile	95% Confidence Interval (±2SE)		75th Percen tile	95th Percen tile	
1	Russian Federation	581	(2.2)	▲	465	540	576	585	626	684
10	England	559	(1.9)	▲	421	508	555	562	613	680
14	United States	549	(3.1)	▲	410	501	543	556	604	666
33	France	511	(2.2)	▲	389	468	507	516	559	617

As you will recall Alexander argued that in Russia and in France,

The quality and power of children's spoken language gain immeasurably from [the episodic lesson structures with their relatively fast interactive and cognitive pace], as one would expect. Further, there is no evidence that the development of children's reading and writing are in any way disadvantaged as a result. On the contrary, the relationship and function of each seem to be better understood. (op cit. p. 565, underline added)

The PIRLS data have some relevance to Alexander's comparisons of the countries. Although in 2016 Russian students performed well in the reading assessments, students in France did not. And in spite of the neglect of oral language England's ranking improved in comparison to the previous PIRLS assessment round in 2011. At the very least these data suggest that whether reading is disadvantaged, or advantaged, as a result of the approach to oral language in the curriculum is a moot point.

As with the critique of the qualitative methodology that we offered above there are issues about the methodology of PIRLS that need to be considered. For example, questions are often raised about the statistical significance of any comparison between two or more countries. The PIRLS reports address this, and in relation to the four countries the relative positions are all statistically significant (IEA, 2016b). In addition, the upward arrows in Table 1 indicate that the countries were all assessed as above the 'centrepoint of the PIRLS scale' which was located in 2001 at 500, the mean of the combined achievement distribution. Another important methodological issue is the validity and reliability of the test to measure reading ability. Comparability across languages is carefully addressed by the designers of PIRLS but many would argue that being a reader in the fullest sense is a culturally specific activity. Another common criticism is that due to the nature of paper-based tests they tend towards low-level short answer questions which are not able to evaluate more sophisticated forms of comprehension. It is true that the test format means that answers are in the main short, but in addition to multiple choice answers they do include the requirements for students to write some answers in open space formats. And the questions that appear later in the tests are designed to assess aspects such as "Process: Interpret and Integrate Ideas and Information". One example of this in the 2016 tests was for students being tested to explain, in the story to be comprehended, the significance of the turn of phrase "at the top of the pecking order" (IEA, 2016c, p.353) in relation to the main character Macy's clever plan to save the hen from predators.

In addition to the quantitative test data, and the extensive methodological publications, there are further sources of data that could be used for comparison of different countries. In each cycle PIRLS/TIMSS country administrators appointed by the international organisers are asked to complete a questionnaire regarding the curriculum arrangements in their jurisdiction (for example, for TIMSS2011, Mullis et al., 2012a, and for PIRLS 2011, Mullis et al., 2012b). Similarly, in each cycle, questionnaires are completed by school leaders and teachers in the stratified random sample of

schools in which pupils take the PIRLS and TIMSS tests (to which we return below).

The national-level data include a core of questions on the nature of the curriculum that have remained constant, or at least very similar, across the years capturing: the existence of a national curriculum; specificity of the curriculum e.g. goals, processes, or materials; the form in which the curriculum is made available; whether the curriculum prescribes the percentage of instructional time devoted to this part of the curriculum; and the way in which the curriculum implementation is evaluated. These concepts link, albeit imperfectly, with theoretical constructs underpinning curricula, as discussed above. This allows for the potential of comparing how these differ between the five countries at any given point in time, along with the extent to which we observe differential changes in these countries over the successive rounds of TIMSS (every three years) and PIRLS (every five years).

It is important to be aware of the imperfect nature of these measures and the way in which they are collected. For example, it is difficult for a national administrator to provide a consistent report for their jurisdiction where policy is highly federal; similarly, country administrators of TIMSS and PIRLS change over time and, even where the organisation does not change, personnel changes could introduce differences in question interpretation. Nevertheless, we argue that national-level comparative data on curriculum arrangements bring some advantages (as well as disadvantages) compared to work that has attempted to consider country-level policy using data collected from school leaders (e.g. OECD, 2011) or from more ad-hoc consideration of national governments' policy documents. Given the complexity of the concepts under consideration, it is a distinct advantage to use this deliberately internationally-comparative data in which a common language (English) is used, aiding the comparability of responses to the questions posed to country administrators. However, we should not overlook the continued presence of some degree of translation-interpretation issues when comparing between countries with differing languages or understandings of specific concepts (which could, of course, differ between countries sharing a language). A noticeable trend in critiques of international comparative work has included questions raised about comparison of countries where different dominant languages are present (e.g. Hilton 2006).

The issue of lack of attention to oral language (particularly in relation to England when compared with the four other countries) is also the case in the information that is available through quantitative international comparisons. Oral language is left out of the scope of PIRLS (as suggested by its name as the "Progress in International Reading Literacy Study), potentially because it is a difficult, time-consuming and costly aspect to assess, compared with the testing of reading for example.

In the absence of specific data about oral language, in order to explore a specific area of the curriculum we have selected reading. Reading as an area in the curriculum does have significant connections with oral language because oral language is the basis for the development of literacy. What is

more, reading for pleasure is an aspect that is individual to the learner but also a concern of teachers and school systems, just as oral language is.

Table x.2. Proportion of PIRLS06 and PIRLS11 countries reporting levels of emphasis on reading for pleasure at each time point

Emphasis on reading for pleasure %	2006	2011
little or no emphasis	9.1	15.2
some emphasis	33.3	45.5
major emphasis	57.6	39.4
Total	100	100

Notes. Data from curriculum questionnaires completed by countries participating in PIRLS2006 and PIRLS2011. N=33. Reporting column percentages.

Table x.2 finds a decrease in the proportion of these countries who report a major emphasis on reading for pleasure in their primary reading curriculum (from more than half to below 40%), along with an increase in the proportion of countries reporting little or no emphasis on this aspect (from under 10% to more than 15%); taken together these suggest a reduction in the emphasis on reading for pleasure in primary curricula across the world (to the extent that it is well represented by the available countries). We may consider the dynamics underlying these aggregate trends with a transition matrix (Table x.3).

Table x.3. Transition matrix of PIRLS06 and PIRLS11 of levels of emphasis on reading for pleasure

Emphasis on reading for pleasure %	2011 little or no emphasis	2011 some emphasis	2011 major emphasis	Total
2006 little or no emphasis	0	100	0	100
2006 some emphasis	18.2	45.5	36.4	100
2006 major emphasis	15.8	36.8	47.4	100
Total	15.2	45.5	39.4	100

Notes. Data from curriculum questionnaires completed by countries participating in PIRLS2006 and PIRLS2011. N=33. Reporting row percentages.

While we reiterate the caveats we emphasize above regarding interpretation of these data, we think this is a meaningful basis for further investigation of trends that may otherwise go unnoticed. For example, none of the countries who said they had “little or no emphasis” on reading for pleasure in their primary curriculum in 2006 still report that this is the case by 2011. This raises the question of whether such a phenomenon might represent a significant change in curriculum development internationally. However, since the group reporting “little or no emphasis” in 2006 only represents three countries (n=33), we should not simply draw this as a conclusion. Instead, it raises a

hypothesis that could be explored with a) subsequent rounds of data, in which we can see what becomes of the countries that have moved into this category in 2011 and/or b) case studies on the dynamics of this shift in policy in these three countries. The question about reading for pleasure also prompts questions about the extent to which reading for pleasure is a focus for classroom interaction in classrooms, research questions that are amendable to qualitative and quantitative methods (e.g. pupil surveys).

The focus on a specific area of the curriculum, the teaching of reading, can be extended to the education system level by looking at another area covered in the questionnaires: the use of inspection. This is linked with performativity, and hence control exerted on schools by such mechanisms as those discussed earlier in this chapter. The changes observed in the proportion of countries reporting that inspection is used to assess implementation of primary reading curricula in the 2006 and 2011 rounds. For the purposes of illustration, we use the full sample available across both rounds, rather than restricting to the five countries discussed above; however, in use of such analyses to address research questions the issue of an appropriate sample of countries would be important to consider.

The cross-sectional percentages reporting that inspection is used to this end in their country are reported in Table x.4, providing illustration of the prevalence of this practice at each time point and suggesting an aggregate increase in the proportion using inspection among this sample. Again, more information on the changes underlying these aggregate shifts may be explored using a transition matrix (Table x.5). This reports the percentage of those who did not use inspection in 2006 who have a) continued not to do so and b) started doing so, along with the percentage of those who did use inspection in 2006 who have c) stopped doing so and d) continued to do so.

Table x.4. Percentage of PIRLS06 and PIRLS11 countries where inspection is used to assess implementation of primary reading curriculum at each time point

%	2006	2011
Inspection not used to assess reading	36.4	27.3
Inspection used to assess reading	63.6	72.7
Total	100	100

Notes. Data from curriculum questionnaires completed by countries participating in PIRLS2006 and PIRLS2011. N=33. Reporting column percentages.

Table x.5. Transition matrix of PIRLS06 and PIRLS11 of inspection use to assess implementation of primary reading curriculum

Inspection use to assess reading	2011		
	2011	2011	2011

curriculum			
%			
2006	Inspection not used	Inspection used	Total
Inspection not used	58.3	41.7	100
Inspection used	9.5	90.5	100
Total	27.3	72.7	100

Notes. Data from curriculum questionnaires completed by countries participating in PIRLS2006 and PIRLS2011. N=33. Reporting row percentages.

This analysis of changes suggests that the overall increase in the proportion of sampled countries using inspection to assess implementation of primary reading curricula is not a one-way street. While some countries that weren't using inspection in 2006 were doing so by 2011 (42%), it is also the case that a proportion of those who were using inspection in 2006 reported that they had stopped doing so by 2011 (just under 10%). While we again emphasise the limitations inherent in these data, differential change of this kind raises hypotheses that education policy may not be best characterised by simple policy convergence (Bieber, 2016) but, rather, may be characterised by more complex dynamics that may be of further interest to explore (Jakobi & Teltemann, 2011).

It is also possible to use these data to explore curriculum implementation. TIMSS (but not PIRLS) asks questions on which areas the curriculum covers, not only at the national-level, but also to teachers of pupils participating in the associated attainment tests. These questions link directly to the areas tested as part of the TIMSS attainment tests. Adopting the definition of curriculum in three elements - an intended curriculum (at the national-level), an implemented curriculum (what is taught in classrooms) and the attained curriculum (what students learn) (Mullis et al., 2009) – the presence of data on these issues at all three levels potentially allows for the exploration of relationships between these levels. We acknowledge, however, that these questions have been less stable over time, which poses significant challenges for analyses using these data (Suter, 2017).

The school- and teacher-level data in TIMSS also provide important details of contextual factors at each of these levels. At school-level these include details of school size, socio-economic context, and management practices. At teacher-level these include details of teacher age, experience, qualification, and job satisfaction. In addition, the TIMSS teacher questionnaire includes reports of the areas of the curriculum they implement, corresponding to the questions asked in the national curriculum questionnaire. These contextual factors would have been relevant to the system level analyses in the Five Cultures study.

In this section, we have highlighted some potential uses of quantitative data and analysis to explore cross-national differences in primary curricula. This

approach does not capture the same richness as in-depth qualitative inquiry discussed above; we do not pretend otherwise. Nevertheless, it can provide opportunities for developing unexplored hypotheses, additional insights, and valuable context (for example from more representative samples and/or wider ranges of countries), than is feasible when conducting in-depth qualitative work alone.

Discussion and conclusions

Most research that engages with curriculum as an object of study is either explicitly located in theories of knowledge, or links implicitly with theories of knowledge through its selection of curriculum elements for focus, for example the differing emphases on oral language versus written language reviewed in this chapter. Our analysis of qualitative and quantitative methods to examine national curricula in primary education suggests that there may be a way forward in relation to Muller's question about knowledge that is both 'constructed' and 'real'. Rapprochement in curriculum studies may be extended through methodologies that seek to constructively use large scale data sets in comparison with in-depth empirical enquiry within countries. With regard to knowledge in the curriculum, powerful or otherwise, rapprochement may also lie in the ways in which both so called powerful and non-powerful aspects might be represented, balanced and enacted as a result of curriculum specifications.

One of the historical features of comparative work that Julliene identified is in need of renewed attention through international comparative work. Aims for education are fundamental to any nation's education system, and as a result require value judgements to be made in the context of degrees of democratic involvement of stakeholders. The question of whether performativity pressures are leading to homogenisation of national curriculum aims is an important one because if found to be the case then it is possible that democratic involvement of countries' citizens in the development of national aims is being replaced by un-critical acceptance that patterns of similar aims in jurisdictions scoring highly in international comparative testing is a sufficient warrant for such aims.

The methodology of qualitative comparisons allows for significant depth to include some of the daily interactions of pupils and teachers as their curricula are enacted. There is also a significant tradition of critical attention to national curriculum texts on the basis of their inherent logic, or lack of logic, and the complex differences between texts and the practice in schools. Yet the problem of generalisation remains a real one for qualitative work, in spite of the significant depth and theorisation that is a feature of the best work.

The methodology of quantitative comparisons allows for significant breadth. This breadth, including large scale testing of pupil achievement, does allow for generalisations on the basis of statistical analyses. However, too often the careful caveats expressed by the authors of these large-scale studies are ignored by politicians as they seek to justify their policies on the basis of

correlations between a country's position in world rankings and selected policies. Inevitably the breadth results in lack of depth in some areas, for example a lack of attention to oral language, and to some of the many cultural, political and historical aspects that are a defining feature of work in the qualitative traditions.

In conclusion, on the basis of our analysis of the methodology of some of the most notable qualitative and quantitative work the limitations of both strongly suggest the need for studies that combine quantitative and qualitative comparative methodology. Our work in this chapter suggests that oral language should be attended to in large scale comparisons. This would represent using the findings from in-depth qualitative work, such as the Five Cultures study, as the basis for building change to large scale quantitative work. It would also be possible to build change on the basis of some of the findings of quantitative comparison by undertaking in-depth enquiry, for example to examine the validity of claims about changes in motivation for reading. These kinds of mixed methods combinations have great potential to add original findings to the considerable findings already established separately through the different traditions. The urgency for this kind of new work on primary education curricula is perhaps most starkly underlined in the worryingly unequal access to high quality primary education, including the lack of evidence-informed exemplary curriculum, pedagogy and assessment, that is a feature of so many of the education systems of the world today.

References

- Alexander, R. (2000). *Culture and Pedagogy: International Comparisons in Primary Education*. Oxford: Blackwell.
- Alexander, R. (2011). Evidence, rhetoric and collateral damage: the problematic pursuit of 'world class' standards. *Cambridge Journal of Education*, 41(3), 265-286.
- Alexander, R., Doddington, C., Gray, J., Hargreaves, L., & Kershner, R. (Eds.). (2010). *The Cambridge Primary Review Research Surveys*. London: Routledge.
- Ball, S. (2012). *Global Education INC.: New Policy Networks and the Neo-Liberal Imaginary*. London: Routledge.
- Bernstein, B. (1971). *Class, Codes and Control. Volume 1: Theoretical Studies towards a Sociology of Language*. London: Routledge and Kegan Paul.
- Bernstein, B. (1996). *Pedagogy, Symbolic Control and Identity*. London: Taylor & Francis.
- Bieber, T. (2016) *Soft Governance, International Organizations and Education Policy Convergence*, London: Palgrave Macmillan.
- Biesta, G. (2014). Pragmatism in the curriculum: bringing knowledge back into the curriculum conversation, but via pragmatism. *The Curriculum Journal*, 25(1), 29-49.
- Carvalho, L. M., & Costa, E. (2015). Seeing education with one's own eyes and through PISA lenses: considerations of the reception of PISA IN European countries. *Discourse: Studies in the Cultural Politics of Education*, 36(5), 638-646.
- Cresswell, J. (2009). *Research design: Qualitative, quantitative, and mixed-methods approaches* (3rd ed.). Los Angeles: SAGE.
- Forestier, K., & Adamson, B. (2017). A critique of PISA and what Jullien's plan might offer. *Compare: A Journal of Comparative and International Education*, 47(3), 359-373.
- Fraser, S. (1964) *Julien's Plan for Comparative Education 18-16-1817*. Teachers College, Columbia University: Bureau of Publications.
- Goswami, U. (2008). *Cognitive Development: The Learning Brain*. Hove: Psychology Press.
- Hilton, M. (2006). Measuring Standards in Primary English: Issues of validity and accountability with respect to PIRLS and National Curriculum test scores. *British Educational Research Journal*, 32(6), 817-837.
- Hopman, S. (2010). *When the battle's lost and won. Some observations concerning "What ever happened to Curriculum Theory"*. University of Stirling. Stirling.
- IEA TIMMS & PIRLS International Study Centre. (2016a). Student Achievement: Distribution of Reading Achievement. Retrieved from <http://timssandpirls.bc.edu/pirls2016/international-results/pirls/student-achievement/pirls-achievement-results/>
- IEA TIMMS & PIRLS International Study Centre. (2016b). Student Achievement: Multiple Comparisons of Average Reading Achievement. Retrieved from <http://timssandpirls.bc.edu/pirls2016/international-results/pirls/student-achievement/multiple-comparisons-of-reading-achievement/>

- IEA TIMSS & PIRLS International Study Centre. (2016c). Appendix H: Restricted Use Passages, Questions, and Scoring Guides. Boston: IEA TIMSS & PIRLS International Study Centre.
- Jakobi, A. & Teltemann, J. (2011) Convergence in education policy? A quantitative analysis of policy change and stability in OECD countries, *Compare: A Journal of Comparative and International Education*, 41:5, 579-595, DOI: 10.1080/03057925.2011.566442
- Jerrim, J., Parker, P., Chmielewski, A. K. & Anders, J. (2016) Private schooling, educational transitions and early labour market outcomes: Evidence from three Anglophone countries. *European Sociological Review*, vol. 32, no. 1, pp. 280-294.
- Lyotard, J.-F. (1984). *The Postmodern Condition: A Report on Knowledge* (G. Bennington & B. Massumi, Trans.). Manchester: Manchester University Press.
- Muller, J. (2016). Knowledge and the Curriculum in the Sociology of Knowledge. In D. Wyse, L. Hayward, & J. Pandya (Eds.), *The SAGE Handbook of Curriculum, Pedagogy and Assessment*. London: SAGE.
- Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., Preuschoff, C. (2009) TIMSS 2011 Assessment Frameworks. Chestnut Hill, MA: TIMSS & PIRLS International Study Centre, Boston College.
- Mullis, I. V. S., Martin, M. O., Minnich, C.A., Stanco, G.M., Arora, A., Centurino, V.A.S., & Castle, C.E. (2012a). TIMSS 2011 Encyclopedia: Education Policy and Curriculum in Mathematics and Science, Volumes 1 and 2. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M.O., Minnich, C.A., Drucker, K.T., & Ragan, M.A. (2012b). PIRLS 2011 Encyclopedia: Education Policy and Curriculum in Reading, Volumes 1 and 2. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- OECD. (2011). School autonomy and accountability: Are they related to student performance? PISA In Focus. Paris, France: Organisation for Economic Cooperation and Development
- Reiss, M., & White, J. (2013). *An Aims-based Curriculum: The significance of human flourishing for schools*. London: IOE Press.
- Spencer, H. (ND [Originally published 1860]). *Education: Intellectual, Moral and Physical*. New York: Hurst & Company.
- Suter, L. (2017) How international studies contributed to educational theory and methods through measurement of opportunity to learn mathematics. *Research in Comparative & International Education*, 12(2), 174-197. DOI: 10.1177/1745499917711549
- Washbrook E., Waldfogel, J., Bradbury, B., Corak, M. and Ghanghro, A. (2012). The development of young children of immigrants in Australia, Canada, the United Kingdom and the United States. *Child Development*, 83 , 1591 – 1607 .
- Wyse, D., Hayward, L., & Pandya, J. (2015). Introduction: Curriculum and its Message Systems: From Crisis to Rapprochement. In D. Wyse, L. Hayward, & J. Pandya (Eds.), *The SAGE Handbook of Curriculum, Pedagogy and Assessment*. London: SAGE.
- Wyse, D., Jones, R., Bradford, H. & Wolpert, M. A. (2018 in-press). *Teaching English, language and literacy* (Fourth Edition). London: Routledge.

- Wyse, D. (2017). *How Writing Works: From the Invention of the Alphabet to the Rise of Social Media*. Cambridge: Cambridge University Press.
- Young, M. (2008). *Bringing Knowledge Back In: From social constructivism to social realism in the sociology of education*. London: Routledge.
- Young, M. (2013). Overcoming the crisis in curriculum theory: a knowledge-based approach. *Journal Of Curriculum Studies*, 45(2), 101-118.
- UNESCO (2017). Sustainable Development Goals. Available from <http://www.un.org/sustainabledevelopment/education/>