

Improving Learning, Fostering the Will to Learn	 Earli
	European Association for Research on Learning and Instruction

The Problematics of International Lesson Structure Comparisons

David Clarke, University of Melbourne, Australia
<d.clarke@unimelb.edu.au>

Paper presented as part of the symposium “Perspectives on International Comparisons of Lesson Structure in Mathematics Classrooms in Germany, Japan, the USA, and Australia” at the 10th Biennial Conference of the European Association for Research on Learning and Instruction, Padova, Italy - August 26 – 30, 2003.

Other papers in the same symposium are:

Mesiti, C. & Clarke, D. “Lesson Patterns in Superficially Similar Cultures: The USA and Australia.”

Jablonka, E. “The Structure of Mathematics Lessons in German Classrooms: Variations on a Theme.”

Shimizu, Y. “Capturing the Structure of Japanese Mathematics Lessons as Embedded in the Teaching Unit.”

The research reported in this paper was supported by grants from the Australian Research Council and the Spencer Foundation (USA).

The Problematics of International Lesson Structure Comparisons

David Clarke, University of Melbourne, Australia

Abstract

How are the classroom practices of different countries most usefully compared if our goal is the improvement of those classroom practices? The aspiration to compare at one level (international) implies an aspiration to typify at another level (national). Both processes, comparison and typification, should be subjected to scrutiny. Lesson structure provides one potential unit of comparative analysis. Lesson structure can be interpreted in three senses:

- (a) At the level of the whole lesson - regularity in the presence and sequence of instructional units of which lessons are composed;*
- (b) At the level of the topic – regularity in the occurrence of lesson elements at points in the instructional sequence associated with a curriculum topic, typically lasting several lessons;*
- (c) At the level of the constituent ‘lesson events’ – regularity in the form and function of types of lesson events from which lessons are constituted.*

This paper addresses the problematics of using Lesson Structure as the basis for international comparisons of classroom practice.

One criterion for the utility of lesson structure as a unit of comparative analysis is its potential adaptability. The international adaptability of a practice is dependent on the degree of consonance between the settings and the beliefs of the originating and adapting cultures. The reporting of international research should anticipate the problems of adaptability by attending to the key issues of voice and variation. “Voice” in this context refers not only to the voices of the participants (teachers and students) in classroom settings (who determine and interpret lesson structure), but also to the voices of the interpreting researchers, whose cultural affiliations inevitably contribute to the form of their analyses (attending to and privileging certain lesson elements over others). Concern with “variation” in this paper, relates to the need in international comparative research not to minimize variation by the simplistic aggregation of data at the level of nation (nationally representative ‘lesson patterns’), by the implicit imposition of a common international curriculum through international testing (attributing national performance to national norms of practice), or by the aspiration to “remove” variation through the identification and advocacy of uniform internationally-applicable best practices (the naïve advocacy of Asian classroom practice and lesson structure), but to document and report variation in classroom practice in a manner that anticipates further variation in the adaptation and application of such research (by focusing on constituent lesson elements rather than on the structure of the lesson as a whole). The study of lesson structure is subject to all these concerns, and can be used as a vehicle to highlight the problematics of international comparative research.

The Assumptions of International Comparative Research

It is important to identify the priorities and assumptions of international comparative research. International comparative research is certainly alive and well, judging by the volume of studies reported in the literature. In Raby's bibliography of comparative and international education studies published in 1999 alone, 937 papers from 110 journals are divided into twenty-four categories (Raby, 2000). Mathematics Education was not one of those categories. Categories such as "Gender" and "Higher Education" suggest dimensions across which it is assumed to be legitimate to undertake international comparative research.

Keitel and Kilpatrick (1999) have problematised the assumptions on which international comparative studies of school mathematics are predicated. In particular, they question the treatment of the mathematics curriculum as unproblematic and the associated assumption that a single test can give comparable measures of curriculum effects across countries. They further suggest that the spectre of an "idealized international curriculum" lies behind even the most sophisticated research designs, including text and document analyses and the use of video to study classroom practice.

A pseudo-consensus has been imposed (primarily by the English-speaking world) across systems so that curriculum can be taken as a constant rather than a variable, and so that the operation of other variables can be examined.

(Keitel & Kilpatrick, 1999, p. 253).

Thorsten (2000) makes a related point about the narratives by which education is related to economic strength and the cultural specificity of these narratives. That is, while the conduct of much of international comparative research in mathematics education seems predicated on assumptions derived from conceptions of a global mathematics curriculum, the interpretation of the results of such research will always be framed by politicians, policy makers and curriculum developers in terms of national rather than international aspirations, values, needs and conditions. The situatedness of such interpretations is not only inevitable, it is eminently sensible – even though the actual interpretations may be driven by political rather than educational motives, by a spirit of competition rather than cooperation, and by principles of elitism rather than equity.

We need to challenge the assumption that international comparative studies in education are necessarily evaluative, much less competitive. The potential utility of international comparative research should not be discounted because of the current preoccupation with competition. The locally produced narratives that are fueled by international comparative research could be about the mutual benefits of sharing good practice and about the adaptive potential of the policies and practices of other educational systems to our own. Those who cite the absurdity of comparing apples with oranges need to consider the benefit of grafting new varieties onto sturdy old stock. The fruitfulness of the product will depend significantly on the (cultural) compatibility of the grafted material and the original stock.

Central to the conduct and use of international comparative research is the position adopted with respect to variation. Depending on the researcher's affiliations, variation is seen either as something that international research aspires to minimize or remove, or as an insurmountable barrier to the utilization of international research. The accommodation of variation in

international research will not be achieved by either the identification of universally applicable ‘big theories’ or ‘best practice’ or by the pessimistic presumption of the incommensurability of international differences in educational policy and practice. International comparative research should document and report variation in educational policy and practice in a manner that anticipates further variation in the adaptation and application of such research. This goal is best achieved if our research designs maximize and optimize the contribution of those voices in which our research is constituted: Not only the voices of the participants (teachers and students) in classroom settings (relevant to our focus on lesson structure), but also to the voices of the interpreting researchers, whose cultural affiliations inevitably contribute to the form of their analyses.

Meaningful International Comparisons

What are the contending bases for international comparative research and how is lesson structure situated within the logic of this research? One of the most widely reported results from studies of international assessment of student achievement such as the Third International Mathematics and Science Study (TIMSS) (Beaton & Robitaille, 1999) has been the high national mean scores for students from ‘Asian’ countries. This appears to have triggered the following (naïve) line of reasoning: If Asian countries are consistently successful on international measures of mathematics performance, then less-successful non-Asian countries would do well to adapt for their use the instructional practices of Asian classrooms. Such a line of reasoning is grounded in four key assumptions: (i) that the term “Asian” identifies a coherent body of practice (of which lesson structure is one component); (ii) that the performances valued in international tests constitute an adequate model of mathematics, appropriate to the needs of the less-successful country; (iii) that differences in mathematical performance are attributable to differences in instructional practice, such as lesson structure (and not to other differences in culture, societal affluence or aspiration, or curriculum); and (iv) that the distinctive instructional practices (norms of lesson structure) of more-successful countries (should these exist) can be meaningfully adapted for use by less-successful countries. Each of these key assumptions can be problematised on a variety of grounds (eg Clarke, 2003; Westbury, 1992).

In his re-analysis of data from the Second International Mathematics Study (SIMS), Bracey (1997) suggested that the differences in mathematics performance found at an international level were replicated in a partitioning of the U. S. sample along cultural or ethnic lines. As a simple illustration of this point: Asian-American students, participating in a school system that has been substantially maligned in the U. S. popular press, perform at a level comparable with their high-performing counterparts in schools in Asian countries. This single illustration suggests that differences on particular measures of mathematical performance are at least as attributable to the cultural affiliation of the students as to the particular school system attended. The significance of such internal cultural variation is lost in the aggregation of performance data for countries as culturally plural as the USA, Australia, or Canada. Such analyses also have implications for societies with a small number of substantial ethnically-distinct communities, such as Malaysia and South Africa.

Berliner reiterated this point in an article in the *Washington Post* (Sunday, January 28, 2001, p. B3). That is, rather than serving an agenda of international competitive comparison, the

results of international achievement testing can be analysed to identify members of a nation who are less well served by the school system than others.

Which America are we talking about? . . . Average scores mislead completely in a country as heterogeneous as ours . . . The TIMSS-R tells us just what is happening. In science, for the items common to both the TIMSS and the TIMSS-R, the scores of white students in the United States were exceeded by only three other nations. But black American school children were beaten by every single nation, and Hispanic kids were beaten by all but two nations. A similar pattern was true of mathematics scores . . . The true message of the TIMSS-R and other international assessments is that the United States will not improve in international standings until our terrible inequalities are fixed.

(Berliner, 2001, B3).

A corollary to this line of reasoning is voiced by Wang (2001) who, in discussing technical concerns with TIMSS, cites Hu (2000, p. 8) as saying, "This study does not break down Americans by race, if they did, Asian Americans would likely score as high as Asians in their home countries, and Whites would rank near top of the European nations." There are several ways to interpret this observation. Berliner's approach seems the most rational and productive: From several perspectives the comparison of national means of student achievement is problematic. Comparisons between sectors of the community within a given country may be more fruitful, within a given state or school system even more so. Such comparisons may at least highlight community groups who are less equal in the benefits they accrue from a school system intended to benefit all students equally. Educational policy can then be framed to address any inequalities.

But what are the implications from the perspective of cultural traditions. The analyses summarized above suggest that the cultural affiliation of the learner (whatever their geographical location) is at least as important as the cultural alignment of the school or school system and certainly should not be simplistically identified with nationality.

The previous remarks are not intended to challenge the premise that school systems enact cultural values. However, they do challenge the simplistic identification of culture with nationality. Once the identification (confusion) of nation with culture has been problematised, then the utility of international comparative research can be considered with greater cultural sensitivity.

Similarity and Difference in International Comparative Research

The central problem of international comparative research in education can be summarized as:

How are the educational systems of different countries most usefully compared if our goal is the improvement of those educational systems?

It is useful to consider briefly the different facets of those educational systems that have formed the basis of international comparative research.

Schmidt, McKnight, Valverde, Houang and Wiley (1997) investigated the mathematics curricula of the “almost 50” countries participating in the Third International Mathematics and Science Study (TIMSS). The documented differences in curricular organisation were extensive. Even within a single country differentiated curricula catered to communities perceived as having different needs. Countries differed in the extent of such differentiation, in the complexity or uniformity of their school systems, and in the distribution of educational decision-making responsibility within those school systems. Given such diversity, the identification of any curricular similarity with regard to mathematics should be seen as significant. And there were significant similarities. There were similarities of topic, if not of curricular location; broad correspondences of grade level and content that became differences if you looked more closely; differences in the range of content addressed at a particular grade level, but which repeated particular developmental sequences where common content was addressed over several grade levels. In another international study of mathematics curricula, the OECD study of thirteen countries’ innovative programs in mathematics, science and technology found that, “Virtually everywhere, the curriculum is becoming more practical” (Atkin & Black, 1997, p. 24). Yet, despite this common trend, the same study found significant differences in the reasons that prompted the new curricula (Atkin & Black, 1996). These interwoven similarities and differences are almost the signature of international comparative research.

Schmidt, McKnight, Valverde, Houang, and Wiley (1997) reported that differences in the characterization of mathematical activity were extreme at the Middle School level; from ‘representing’ situations mathematically, ‘generalizing’ and ‘justifying’ to ‘recalling mathematical objects and properties’ and ‘performing routine procedures.’ Despite the apparent diversity, it was the latter two expectations that were emphasised in the curricula studied. Given the documented diversity, it is the occurrence of similarity that requires explanation. Some curricular similarities may be the heritage of a colonial past. Others may be the result of more recent cultural imperialism or simply good international marketing.

In attempting to tease out the patterns of institutional structure and policy evident in international comparative research (particularly in the work of LeTendre, Baker, Akiba, Goesling, and Wiseman, 2001), Anderson-Levitt (2002) noted the “significant national differences in teacher gender, degree of specialization in math, amount of planning time, and duties outside class” (p. 19). But these differences co-exist with similarities in school organization, classroom organization, and curriculum content. Anderson-Levitt (2002, p. 20) juxtaposed the statement by LeTendre et al. that “Japanese, German and U.S. teachers all appear to be working from a very similar ‘cultural script’” (2001, p. 9) with the conclusions of Stigler and Hiebert (1999) that U.S. and Japanese teachers use different cultural scripts for running lessons. The apparent conflict is usefully (if partially) resolved by noting with Anderson, Ryan and Shapiro (1989) that both U.S. and Japanese teachers draw on the same small repertoire of “whole-class, lecture-recitation and seatwork lessons conducted by one teacher with a group of children isolated in a classroom” (Anderson-Levitt, 2002, p.21), but they utilise their options within this repertoire differently.

LeTendre, Baker, Akiba, Goesling and Wiseman (2001) claim that “Policy debates in the U.S. are increasingly informed by use of internationally generated, comparative data” (p.3). LeTendre and his colleagues go on to argue that criticisms of international comparative research on the basis of “culture clash” ignore international isomorphisms at the level of

institutions (particularly schools). LeTendre et al. report yet another interweaving of similarity and difference.

We find some differences in how teachers' work is organised, but similarities in teachers' belief patterns. We find that core teaching practices and teacher beliefs show little national variation, but that other aspects of teachers' work (e.g., non-instructional duties) do show variation.

(LeTendre, Baker, Akiba, Goesling & Wiseman, 2001, p. 3)

These differences and the similarities are interconnected and interdependent and it is likely that policy and practice are best informed by research that examines the nature of the interconnection of specific similarities and differences, rather than simply the frequency of their occurrence. So, where does lesson structure fit into this requirement for interconnection? Not least, in the need to recognize classrooms as sites for the collaborative enactment of the practices and meanings of classroom participants. Lesson structure from this perspective is both a constraint on possible practice and a consequence of that practice.

Focusing on Classroom Practice and the Unit of Comparative Analysis

It seems reasonable to focus attention on the practices of classrooms as the most evident institutionalized means by which the policies of a nation's educational system are put into effect. Specifically, the classroom seems a sensible place to look for explanations and consequences of the differences and similarities identified in international comparative studies of curriculum, teaching practice, and student achievement.

Within the specific focus of Classroom Practice, the central problem of international comparative research translates into:

Are there nationally-specific characteristics of classroom practice?

How best might the practices of classrooms be compared internationally if our purpose is to inform those practices?

NB. The second question remains both valid and important, whether the answer to the first question is "Yes" or "No."

The curriculum is the embodiment of the aspirations of the school system. To a significant extent, the teacher is the agent of the system by whose actions the curriculum is put into effect. Teachers, however, interpret the curriculum in idiosyncratic fashion, within the constraints and affordances of both system and culture. Both the curriculum and the teacher have been the focus of recent international comparative study. Among the studies of curriculum and teaching practice, we can lose sight of the student. Thorsten makes this point beautifully.

What is absent from nearly all the rhetoric and variables of TIMSS pointing to the future needs of the global economy is indeed this human side: the notion that students themselves are agents. TIMSS makes students from 41 countries

into passive objects of 41 bureaucratic gazes, all linked to the seduction of one global economic curriculum.

(Thorsten, 2000, p. 71).

As educational research has increasingly drawn our attention to the importance of the social processes whereby competence is constructed and in which competence is constituted (for both teaching and learning). The agency of the student, the nature of learner practice, and the cultural specificity of that agency and that practice must be accommodated within our research designs.

Studying Lesson Structure

The analysis of video data collected in the video component of TIMSS (Stigler and Hiebert, 1999) centred on the teacher's adherence to a culturally-based "script." Central to the identification of these cultural scripts for teaching were the "lesson patterns" reported by Stigler and Hiebert for Germany, Japan and the USA, and the contention that teaching in each of the three countries could be described by a "simple, common pattern" (Stigler & Hiebert, 1999, p. 82).

By contrast, the Learner's Perspective Study (<http://www.edfac.unimelb.edu.au/DSME/lps/>) analysed sequences of ten lessons, documented using three video cameras, and supplemented by the reconstructive accounts of classroom participants obtained in post-lesson video-stimulated interviews. This methodological approach offers an informative complement to the survey-style approach of the TIMSS video study. A research design predicated on a nationally representative sampling of individual lessons, as in TIMSS, inevitably reports a statistically-based characterization of the representative lesson. A more fine-grained study of sequences of ten lessons, informed by the reconstructive accounts of the participants, has the potential to address:

- Consistency of lesson structure over a ten lesson sequence
- Degree of variation in lesson structure in the practices of competent teachers
- The extent to which any such variation is linked to the location of the lesson in the instructional sequence and to the teacher's instructional intentions
- Student awareness of the structure of the lesson and how this is related to their perception of significant educational moments in the lesson and to their subsequent learning.

Lesson structure can be interpreted in three senses:

- (a) At the level of the whole lesson - regularity in the presence and sequence of instructional units of which lessons are composed;
- (b) At the level of the topic – regularity in the occurrence of lesson elements at points in the instructional sequence associated with a curriculum topic, typically lasting several lessons;
- (c) At the level of the constituent lesson events – regularity in the form and function of types of lesson events from which lessons are constituted.

In terms of international comparison, it is useful to consider which of these three forms of lesson structure are likely to prove useful as units of comparative analysis. Given the

distinction, drawn earlier, between the processes of national typification and international comparison, it is important to recognize that the most appropriate unit for national typification may not prove useful for international comparison. In terms of national typification, we need to address the question “Is a nation’s or a culture’s classroom practice, most usefully characterized at the level of the whole lesson, in the manner in which a topic is constructed, delivered and experienced, or in the form and function of the specific activities from which lessons are composed?” The same three alternatives are available for the purposes of international comparison.

There are two quite distinct methodological alternatives:

Alternative 1.

If two groups of objects are to be compared then one approach is to consider these two questions:

Difference – “What is the characteristic about which the *comparison* is to be made?”

Similarity – “How might each group of objects be separately *typified* with respect to that characteristic?”

The international comparison of national norms of student achievement could be described as conforming to this approach.

The order in which these two questions are posed is a major methodological signature.

Alternative 2.

If two groups of objects are to be compared, consider these two questions:

Similarity – “Which characteristics appear to *typify* this collection of objects?”

Difference – “What *comparisons* can be made between these two groups of objects using the identified characteristics?”

Posing the questions as in Alternative 2 reduces the danger of constraining the data to a predetermined structure, but may lead to the typification of the two groups by different emergent characteristics, restricting the common bases on which comparison of the two groups might be made. It should be noted also that Alternative 2 assumes a domain within which comparison is sought, such as classroom practice or curricular policy.

In terms of lesson structure, it might be that for one nation or culture there is no nationally characteristic structure to the lesson as a whole, but that particular types of idiosyncratic lesson events offer the most appropriate typification. For another nation or culture, there could be a high degree of regularity to the composition of lessons, or in the sequencing of particular types of instructional activity in the delivery of a topic. Such differences in the form of typification provide a basis for international comparison that reflects something more essential to each than the identification (imposition) of the same structural level as the basis for the comparison. The methodological choice of Alternative 1 makes the basis for comparison a matter of prescription based on either theory or on the prevailing educational priorities of the country conducting the study. Choice of Alternative 2 makes the identification of possible bases for comparison an empirical result of the research.

Issues of Authorship and Voice

In an international comparative study, any evaluative aspect is reflective of the cultural authorship of the study. If we are to make judgements of merit, whether they are about student achievement or classroom practice, we can only do so from the position of the authoring culture. The efficacy of a practice can only be judged to the extent that it achieves a specified goal. The most obvious goal against which to assess the efficacy of a practice is the goal of the individual or school system engaged in the practice. For the purposes of international comparative research, however, it is legitimate for someone outside the system being studied to evaluate a practice relative to their own goals – provided that this distinction is made explicit. For example, a researcher evaluating a particular curriculum structure would draw different conclusions regarding the efficacy of the structure if the evaluating researcher were assessing its potential utility for a school system in which most students only completed tenth grade compared with one for which twelfth-grade completion were the norm. This is only to say that the report of an international comparative study need not be evaluative, but the readership of such a report may engage in evaluation of the report in relation to their own goals, school system and culture.

Whether an international study implements an evaluative agenda or documents the detail needed for the constructive juxtaposition of policy or practice, the issue of cultural authorship is equally relevant. The design of international comparative studies must implement collaborative processes through which educational, philosophical and cultural positions are given voice in the interpretation of data and the reporting of the research. The OECD study of innovative programs in mathematics, science and technology education went some way towards addressing this issue: “A nine-member writing team prepared the final cross-national report (Atkin & Black, 1996). Almost all the countries published their own case studies in the home language for internal distribution” (Atkin & Black, 1997, p. 23). International collaborative studies can implement protocols requiring that the interpretations of data to be included in published reports must be validated by the member researchers from the country providing the data (as is the case for the Learner’s Perspective Study). This form of “member check” parallels established practice among researchers using high inference qualitative data of seeking confirmation of data interpretations from the respondents in primary research studies. Implementation of the same regime at the level of participating research groups in international comparative studies requires the careful negotiation of a range of methodological and ethical issues such as the minimisation of misrepresentation through misinterpretation and the scrupulous monitoring of the possibility of disadvantage to a participating school system through the publication of a report. Such negotiated cross-cultural authorship offers substantial advantages in interpretive richness and cultural validity that outweigh any hypothetical restrictions on freedom of publication.

The other aspect of cultural authorship relevant here is the issue of representation and voice. In commenting on the proliferation of OECD-initiated international comparative research projects, Cohen characterised the OECD as “a club of 29 of the world’s richest countries” (Cohen, 1998, p. 4). Even when less affluent countries participate in international studies, it is frequently as the objects of investigation rather than as partners in the research. Research is conducted from a ‘Western’ perspective and evaluates the practices it studies by ‘Western’ criteria.

This research remains largely bounded by the Western conception of (teacher-centred) pedagogical practice and by implicit social rules pertaining to authority and social participation.

(Fuller & Clarke, 1994, pp. 143-144).

Among the volumes of text prompted by recent international comparative studies such as TIMSS, no country it seems has been as prolific in generating papers as the USA. These papers warrant a separate analysis of their own, for in them one finds the narratives of nationhood (see Thorsten, 2000). We find such narratives in the media reporting of other large-scale international comparative studies such as PISA (the Programme for International Student Assessment; website: www.pisa.oecd.org). As was noted at the commencement of this paper, the situatedness of these narratives is inevitable and even illuminating. But their situated character should be acknowledged and allowed for in the reading of any report of international comparative research. It might even be the focus of such research. In this discussion, however, cultural authorship highlights the dual significance of voice and variation.

Mirrors Not Blueprints: Adaptive Potential

Globalisation seeks to minimise international differences (whether by consensus or imposition) whereas internationalisation seeks to celebrate both the similarities and the differences and to learn from them. This difference can be illustrated by comparing the goal of aspiring to standardize instructional practice in mathematics classrooms internationally and the goal of aspiring to optimize local practice through critical reflection stimulated by consideration of best practice elsewhere.

Of all the approaches to international comparative research, it is the video studies that we might expect to have the greatest potential to inform classroom practice. Stigler, Gallimore and Hiebert (2000) offer two arguments to support the utility of classroom video studies. Each of these will be discussed separately.

Perhaps the most obvious reason to study classrooms across cultures is that the effectiveness of schooling, as measured by academic achievement, differs across cultures If cross-national achievement differences are tied to cultural variations in teaching, we may discover ways of teaching that work better than the ones our society routinely employs. This would allow us to take advantage of the experience of others all over the world who share similar goals.

(Stigler, Gallimore, & Hiebert, 2000, p. 87)

The crucial phrase here is the last one: “who share similar goals.” If the essential criterion by which we evaluate the effectiveness of a country’s classroom practices is academic achievement, then the value which we attach to the documented practices will depend on the correspondence between the means by which that achievement is measured and the goals and values of the school system to which we might adapt the documented practices. Also central to the passage above is the proposition that “cross-national achievement differences are tied to cultural variations in teaching.” This proposition provides the principle assumption on which international comparative studies of teaching practice are predicated, yet it is fair to say that

this proposition has yet to be validated by research. The implicit identification of culture with nation, already problematised, does not strengthen the proposition. Finally, the connection between specific instructional actions by teachers and national measures of academic achievement is a tenuous one at best. With what confidence can specific teacher practices be linked causally to aggregated measures of national performance? The connection between classroom practice and consequent learning could be made far more convincingly if the data on learning outcomes and on classroom practice were collected in the same setting and in relation to the same individuals.

There is another more subtle reason for studying teaching across cultures. Teaching is a cultural activity. Because cultural activities vary little within a society, they are often transparent and unnoticed . . . Cross-cultural comparison is a powerful way to unveil unnoticed but ubiquitous practices . . . Comparative research invites reexamination of the things “taken for granted” in our teaching, as well as suggesting new approaches that never evolved in our own society.

(Stigler, Gallimore, & Hiebert, 2000, pp. 87-88).

This ‘more subtle’ argument is far less problematic than the first. Watanabe (2001) quotes White (1987) as writing “we should hold Japan up as a mirror, not as a blueprint”. This powerful and appealing metaphor can serve as a general characterisation of one of the major uses of international comparative studies of classroom practice. Both the White quote and the passage above place the agency for the interpretation and adaptation of any documented practice with the person looking in the mirror. There is no invocation of absolute best practice – the judgement is a relativist one, and an instructional activity with a high degree of efficacy in Hong Kong may retain little effectiveness when employed in a Swedish classroom, where different cultural values inform and frame the actions of all classroom participants. Most importantly, we are encouraged to study Japanese (or South African or German) classrooms not solely for the purposes of mimicking their practices but for their capacity to support us in our reflection on our own practice. The mutuality of the potential benefit provides further motivation for such research. These arguments apply to the study of lesson structure as one aspect of classroom practice.

The Way Forward: Learning from Similarity and Difference

International comparative research is open to misuse in at least three ways: (i) Through the *imposition* on participating countries of a global curriculum against which their performance will be judged; (ii) Through the *appropriation* of the research agenda by those countries most responsible for the conduct of the study, the design of the instruments, and the dissemination of the findings; and, (iii) Through the *exploitation* of the results of such studies to disenfranchise communities, school systems, or the teaching profession through the implicit denigration of curricula or teaching practices that were never designed to achieve the goals of the global curriculum on which such studies appear predicated. Each of these three points would be cause for significant concern if it were shown that research agency resided exclusively with a particular cultural perspective (for example, Western or East-Asian).

The earlier discussion has already posited “adaptive potential” as a key consideration by which researchers and educators from one school system consider the relevance and utility of the policies and practices of another. Hatano and Inagaki (1998) remind us that the adaptation of pedagogical practice requires consideration of both the practicality of technical implementation and the extent to which the beliefs underlying the adapted practice are in harmony with local cultural values.

If, as I am arguing, the adaptability of a policy or a practice is dependent on the degree of consonance between the settings and the beliefs of the originating and adapting cultures, then it is incumbent upon the researcher to provide details of those settings and beliefs. This detail is also essential if the research is to support our interrogation of our own practice. The level of detail required is considerable, since the researcher cannot anticipate which aspects of the documented setting and its associated practices and beliefs will resonate or jar with the reader’s situation.

If a key criterion for the consideration of a practice as valuable is the learning it promotes, then our research designs must afford plausible connection between a particular classroom practice (eg lesson structure) and consequent learning outcomes. Thorough reporting of the means by which those learning outcomes are documented will allow the reader of the research to identify the extent to which the learning outcomes reflect the goals and values of the reader’s curriculum, school system and culture. The utility of international comparative testing rests with the reader’s endorsement of the test instrument as the legitimate operationalisation of valued (and locally relevant) educational goals.

International comparative research must be undertaken on a basis of mutual benefit to all participants. The fundamental reflexivity that is embodied in the metaphor of the mirror rather than the blueprint should underlie the function of all such studies. We must guard against the cultural imperialism of an implicit global curriculum and, instead, stress the centrality of local interpretation of all findings. Adoption of such a relativist approach avoids the competitive dichotomisation of Eastern and Western cultural traditions. Research that is predicated on an anticipation of mutual benefit should avoid the pitfalls of appropriation, imposition and exploitation discussed above. The detailed collaborative study of international policy and practice in mathematics education, and of the products of that policy and practice, should be undertaken in anticipation of insights into the novel, interesting and adaptable practices employed in other school systems of whatever cultural persuasion, and of insights into the strange, invisible, and unquestioned routines and rituals of our own school system and our own classrooms.

Research into international differences and similarities in student mathematical performance (such as TIMSS and PISA) has limited utility, except as a form of national report card, unless it is accompanied by data that suggest cultural, societal, or instructional variation that might be used to explain such differences and similarities and then to promote improved mathematical learning and associated performance. One of the difficulties in constructing such explanations is the interdependent character of cultural and social practice, the cultural specificity of instructional practice, and the social mediation of cultural values and curriculum emphases in schools and classrooms.

Case studies offer an alternative to the dangers inherent in survey research of superficiality and indiscriminant aggregation. There is some acceptance of the idea that a detailed case study of good classroom practice in one country might usefully inform classroom practice in another country, if only by foregrounding the assumptions underlying the practices in both and catalysing a review of practice by questioning routines and systems that had never previously been questioned. It is not obvious to everyone, however, that comparative case studies can inform policy. LeTendre, Baker, Akiba, Goesling, and Wiseman (2002) argue that they can.

Studies such as Spindler (1987), Tobin, Wu and Davidson (1989), or Shimahara and Sakai (1995) provide powerful insights into the way teacher practice and belief is shaped within different national contexts. Comparative studies that provide this level of ethnographic detail hold much potential for educational policy.

(LeTendre, Baker, Akiba, Goesling, & Wiseman, 2002, p. 23).

For these authors, it is the detail of practice and belief and their interrelationship that holds the promise of informing policy, rather than the documentation of national levels of student achievement, nationally prevalent teaching actions or nationally typical teacher beliefs. Put bluntly, it is process we need to study and adapt not product – whether it is the teacher education process that engendered the beliefs or the processes of the classroom that fostered the achievements. LeTendre et al. would have us believe that these things are best studied in situ over time rather than in snapshots taken from the moon.

From the studies that have been done, we have every reason to believe that it is in these interrelationships that the character and function of culture will most clearly emerge: In the teacher practice that mediates between curriculum content and the student, through the actions and the lesson structure that constitute the enactment of that curriculum in the classroom, together with the beliefs and expectations on which the student's participation is predicated, culminating in the learning of which student achievement is simply the most evident socially-constructed and culturally-mediated correlate. Culture is not outside these things. It is in the combination of these and other elements that culture itself is constituted. Nor is culture a synonym for nationality. As several studies have shown, the culture of the classroom can be constructed differently within a particular country or school system. There are, however, cultural values and beliefs that frame each country's educational endeavours. International comparative research must do more than document cultural differences, it must accommodate them. The Learner's Perspective Study that generated the analyses reported in the other papers in this symposium was designed to provide this accommodation. It is quite possible that lesson structure, at all three levels identified in this paper, may provide one of the more interesting illustrations of the embodiment of culturally-grounded beliefs and values in educational practice. This paper has attempted to outline some of the concerns that need to be addressed by international comparative research, whether into lesson structure or any other aspect of educational policy or practice.

References

- Anderson, L. W., Ryan, D., & Shapiro, B. (1989). *The IEA classroom environment study*. NY: Pergamon.
- Anderson-Levitt, K. M. (2002). Teaching Culture as National and Transnational: A Response to Teachers' Work. *Educational Researcher* 31(3), 19-21.
- Atkin, J. M. & Black, P. (1996). *Changing the Subject: Innovations in Science, Mathematics, and Technology Education*. London and Paris: Routledge and the OECD (cited in Atkin and Black, 1997).
- Atkin, J. M. & Black, P. (1997). Policy Perils of International Comparisons. *Phi Delta Kappan* 79(1), 22-28.
- Beaton, A. E. & Robitaille, D. F. (1999). An Overview of the Third International Mathematics and Science Study. Chapter 3 in G. Kaiser, E. Luna & I. Huntley (Eds) *International Comparisons in Mathematics Education*, London: Falmer Press, 19-29.
- Bracey, G. W. (1997). On comparing the incomparable: A response to Baker and Stedman. *Educational Researcher*, 26(4), 19-26.
- Bracey, G. W. (1998). TIMSS, Rhymes with 'Dims,' As in 'Witted.' *Phi Delta Kappan*, vol. 79, 686-687.
- Clarke, D.J. (2003). International Comparative Studies in Mathematics Education. Chapter 5 in A.J. Bishop, M.A. Clements, C. Keitel, J. Kilpatrick, and F.K.S. Leung (Eds.) *Second International Handbook of Mathematics Education*, Dordrecht: Kluwer, 145-186.
- Cohen, D. (1998). World League Tables: What's the Score? *Principal Matters*, 10(1), pp. 3-7.
- Fuller, B. & Clarke, P. (1994). Raising School Effects While Ignoring Culture? Local Conditions and the Influence of Classroom Tools, Rules, and Pedagogy. *Review of Educational Research*, 64(1), 119-157.
- Hatano, G. & Inagaki, K. (1998). Cultural Contexts of Schooling Revisited: A Review of The Learning Gap from a Cultural Psychology Perspective. Chapter 3 in S. G. Paris & H. M. Wellman (Eds.) *Global Prospects for Education: Development, Culture and Schooling*. Washington, D.C.: American Psychological Association, 79-104.
- Hess, R. D. & Azuma, H. (1991). Cultural Support for Schooling: Contrasts between Japan and the United States. *Educational Researcher*, 20(9), 2-8, 12.
- Hu, A. (2000, December). TIMSS: Arthur Hu's index. [On line] Available: <http://www.leconsulting.com/arthurhu/index/timss.htm>. (December 30, 2000). Cited in Wang (2001).
- Keitel, C. & Kilpatrick, J. (1999). The Rationality and Irrationality of International Comparative Studies. Chapter 16 in G. Kaiser, E. Luna, & I. Huntley (Eds.) *International Comparisons in Mathematics Education*. London: Falmer Press, pp. 241-256.
- LeTendre, G., Baker, D., Akiba, M., Goesling, B., & Wiseman, A. (2001). Teachers' Work: Institutional Isomorphism and Cultural Variation in the US, Germany, and Japan. *Educational Researcher* 30(6), 3-15.
- Raby, R. L. (2000). Comparative and International Education: A Bibliography (1999). *Comparative Education Review*, 44(3), 381-419.
- Schmidt, W. H., McKnight, C. C., Valverde, G. A, Houang, R. T., & Wiley, D. E. (1997). *Many Visions, Many Aims Volume 1: A Cross-National Investigation of Curricular Intentions in School Mathematics*. Dordrecht: Kluwer.
- Stigler, J., Gallimore, R., & Hiebert, J. (2000). Using Video Surveys to Compare Classrooms and Teaching Across Cultures. *Educational Psychologist*, 35(2), 87-100.

- Stigler, J. & Hiebert, J. (1999). *The Teaching Gap*. New York: Simon & Schuster.
- Suri, H. (2002). Essential Features of Methodologically Inclusive Research Syntheses. Paper presented as a part of the symposium, *Primary Research, Secondary Research and Research Synthesis – Theory, Value and Linkage*, at the Annual Meeting of the American Educational Research Association, New Orleans, April 1-5, 2002.
- Suri, H. (in preparation). Methodologically Inclusive Research Synthesis. Doctoral thesis to be submitted to the University of Melbourne, Australia.
- Thorsten, M. (2000). Once upon a TIMSS: American and Japanese Narrations of the Third International Mathematics and Science Study. *Education and Society* 18(3), 45-76.
- Wang, J. (2001). TIMSS Primary and Middle School Data: Some Technical Concerns. *Educational Researcher* 30(6), 17-21.
- Watanabe, T. (2001). Content and Organization of Teacher's Manuals: An Analysis of Japanese Elementary Mathematics Teacher's Manuals. *School Science and Mathematics*, 101(4), 194-201.
- White, M. (1987). *The Japanese educational challenge: A commitment to children*. New York: The Free Press. Cited in Watanabe (2001).