

One pattern – various realizations: The TIMSS lessons in light of a theory of classroom¹

Peter Menck

University of Siegen, Germany

Whenever the results of a new TIMSS-like study are published they cause great concern in one country and the feeling of satisfaction in another. It is not really the outcomes as such that stir the discussion on the respective national educational systems but just the comparison of the overall scores: It is not really the outcomes as such that stir the discussion; it is rather the respective nation's educational system that is questioned. In the first comparative studies the educational system was seen as an independent variable with the average learning outcomes as the dependent one. So the system as a whole was presumed to cause differences between nations. But the system and its formal organization does not explain much as it is embedded in a comprehensive national culture and is composed of a number of components. One of these features, probably the most important one, is the *classroom culture*. From this point of view, the problem of the TIMSS study may be seen as an implicit hypothesis that underlies the first *TIMSS Videotape Classroom Study* from 1995: Different and particularly nation-specific classroom cultures may cause differences in average test scores.

1 The question

What does *classroom culture* mean? One way to get hold of this concept is the classroom itself. Thus a number of video studies have been carried out with piles of classroom recordings as a result. Whenever a more detailed aspect is looked into e. g. classroom as a whole, or classroom management or even the educational system, *classroom recordings* seem to be the method to choose. This is what led James W. Stigler, a psychologist, to conduct his famous study.

Beginning in 1994, the TIMSS Videotape Classroom Study gathered over 250 videotaped recordings of classroom instruction from national samples of classrooms in Germany, Japan, and the United States. The purpose was to provide a rich source

¹ An earlier version of this paper was presented at the 2001 AERA meeting in Seattle.

of information on what goes on inside eighth-grade mathematics classes in these nations [...]. This study provides a wealth pool of classroom-based video data for research efforts aimed at further understanding of the complexities of classroom practices from national and international perspectives (cf. Stigler et al., 1999).

When I developed my argument in 2000, there were no more than six video-tape recordings of lessons available – and those were of rather bad quality (Stigler et al., 1999). But from Stigler's introduction to that collection I presumed that the data of the study as a whole couldn't have been much better than the sample we had at hand. The recordings were restricted to one single lesson per each participating teacher; the audio transcriptions are just a summary of what had happened in the respective classrooms and only poor information is provided.

I just mention in passing that a lot of research on verbal transcriptions of audiotaped lessons was done in the 1960s and 1970s but it looks as if this has been forgotten and in fact done in vain. The small data basis and its deficiencies did not bother me then as it does not now. In the meantime, numerous video studies have been carried out, and tapes of much better quality including even those old ones are sufficient for the argument I'm going to develop, namely the TIMSS question as I may call it. But may recordings as such, even those of better quality, provide appropriate data to test the far reaching TIMSS Video hypothesis? Can we, and if so, how can we infer different test scores from either a handful or even several hundreds of videotaped lessons? Even in case we could, what kind of a theoretical framework would allow us to at least loosely connect classroom patterns on the one hand with learning outcomes on the other? Furthermore, and that is my point here, when we intend nationwide comparisons, we need something like a *third*, a *tertium comparationis*, as formal logic puts it: What kind of reference differences in learning outcomes may be traced back to when comparing classrooms and classroom work?

2 The concept of “classroom”

When speaking of *different* classrooms we obviously have a *concept of classroom* in mind – an average classroom, a classroom as such, a classroom as it should be, or the like. Everybody has their concepts of classrooms we can see when we look at them or into them. But can we be sure that these *everybody's* concepts coincide? So that the result may serve as the *third* we're

looking for? I am afraid we would find out a wide variety of images (rather than concepts in the strict sense) which overlap in some aspects and differ in others. If that was the case – and I am convinced it is – could we look for the smallest commonly shared set of convictions and define that set as the *third* we need for a comparison?

For example, could we observe eighth grade math classrooms in Japan, the US, Germany, or even in all of the 30 countries involved in TIMSS and expect that the result will be something like a universal *concept of classroom*? First of all, this method would be particularly unsuitable for judging differences. Although it would not pose any problem around the features all classrooms have in common, what about those we observe in classroom A or even in country X but *not* in the other ones? The empirical way to find a suitable *third* would be quite a journey with a lot of uncertainties. All these uncertainties in mind, I'm going to suggest a *theoretical* rather than an *empirical* approach in the following. That is to say, I am looking for a *theory of classroom* in the context of which we may discuss the problem of comparisons.

3 The logic of a classroom process

To start with, there are the beginnings of two of lessons from our sample:

T: ... indicates that they will learn a 4th formula... [GG]²

T: ... our study today will use this as a foundation... [JG]

The teachers bring up what they are going to deal with. That means that the students do *not know yet*, and they are promised that they *will know* at the end of the lesson. – At the end we observe a corresponding statement such as e. g.:

T: ... what we've just done and what we've learned... [GG]

The teacher – or sometimes a student – is *summarizing the results* or *the objectives*, as seen in the subtitles of the videos. The students are or at least should be able to do what the classroom work is aimed at – moreover they

² My quotes in the following are from the subtitles in the recordings mentioned above. The first letter refers to the respective countries (G: Germany; J: Japan; U: United States); the second one to the disciplines (G: geometry; A: arithmetic).

know or at least can know that they are able to do. From these observations I derive a *basic assumption*: It is the difference of not yet knowing and knowing which constitutes the process between the beginning and the end of the lesson as a pedagogical or didactical process.

Before exploring the logic of this didactical process I am going to explain my assumption in more detail and to put it in a set of what I call *axioms of a theory of classroom*. By axiom I mean a presupposition which is agreed upon in a certain (scientific) community and needs neither proof nor explication.

- The axiom of culture: In every society there is culture which makes survival and social co-existence possible; this culture is a complement to what nature provides.
- The axiom of tradition: In every society culture is passed on, i.e. culture is transmitted from those who have acquired it to those who have not.
- The axiom of institution: The transmission of culture within a society is institutionalized.
- The axiom of generations: In every society there is an older generation of those who are full members of this society, and a younger generation of those who are not yet full members of the society.
- The axiom of a minimum: The older generation is responsible for the passing on of a cultural minimum.

In short and by the way of a formal definition: *Classroom work is the passing on of a minimum of culture from (and by) the older generation to the younger one.*

No matter whether it is a set of axioms or a classical definition, whenever one of the characteristic features is lacking, we would not speak of *classroom*. On the other hand, these axioms are sufficient to circumscribe what classroom is, or to put it more precisely, classroom as it is seen in the scientific community Stigler and his followers refer to.³

Now I am going to reconstruct the *logic of the classroom process*, i. e. the process between the state of not yet knowing on the one hand and knowing on the other.

³ Although I am convinced that my axioms apply in every human society, I restrict my argument to the discourse community defined by the education issues brought up in the OECD countries.

Knowing serves as an abbreviation here. The term stands for all the abilities the younger generation is to acquire during their school education (according to axiom 2).

1. The students do not yet know the matter they have to learn. Furthermore, at a compulsory school we cannot assume that they are aware of this “not yet”. So first of all, the teacher has to organize the *willingness to work* and to work methodically in particular. This is the pedagogical version of what is well-known as the concept of *motivation* in psychology (in education we should speak of *motivating* instead).
2. The awareness of not yet knowing is usually taken as a basis of a *task* or of a *problem to be solved*. The assumption that underlies a task in classroom is this: Students who are able to solve the problem worked on in classroom will have the competence to solve similar problems on their own (which they may demonstrate in tests like the TIMSS tests).
3. *Tools* must be *provided*. What do we know already and what can we make use of? Media have to be at hand and their use must be explained: e. g. books; knowing persons such as the teacher may help during the work.
4. Next is the *work on the problem* in question which comes to its logical end with one or more *solutions*. A crucial trait of the concept of a teacher is derived from this: There must be at least one person who knows at least one way leading to at least one solution. Whoever that may be – it is up to a teacher to ratify the solution(s). By contrast to medieval classrooms, in modern ones the work pattern is customary no matter how it is organized in detail. That is why next step is obligatory:
5. It has to be guaranteed that *every single student* is able to solve the problem – since every student has the right – and is obligated – to acquire the ability in question, since the older generation and the society in general has the right to expect that knowing subjects are graduating from their schools. For centuries this step has been organized as the *application* i. e. the solution of similar problems, be it in a formal testing situation or in assessment centers.

The classroom process as such is – logically speaking – finished when the teacher, and consequently the older generation, can be assured that all students are able to solve the problem in question and can be considered to have the cognitive, affective, and psychomotor abilities or competences which they are supposed to have according to the respective curriculum.

Those five steps establish the pattern of what I call the *classroom process*; in the tradition of Johann Friedrich Herbart we speak of its *articulation*. As a matter of fact it does not need to be five steps. In the German *Didaktik tradition*, more detailed *models* of the process can be found. And as we know from the history of education there has been a tripartite model: *introductio – explicatio – applicatio*.

My argument is that this pattern, this logic of the classroom process, may serve as our *tertium comparationis*, as *the third of a comparative evaluation of different classrooms* in view of the outcome of classroom work. Apparent differences in the pattern's realization have to be explained by taking culture-specific variables into consideration. By *culture* I mean different inter-national as well as intra-national cultures (e. g. the New Math or reform pedagogy in general vs. traditional mathematics or classroom management).

Moreover, my argument is that there is *one pattern only*. And this argument derives from the fact that in all cases it is *classroom* and *classroom work*. In this I do not at all agree with Stigler who, referring to different national cultures, speaks of different *scripts*. To put my argument in his terms: it is one script, what differs are the realizations in the classroom practice.

4 Comparisons

How could a comparison of our different classrooms work though? We simply have to apply our pattern to the respective lessons. But before doing so we should take into consideration that the process as a whole does not necessarily fit in the unit of time usually dedicated to one lesson, i.e. 45 or 50 minutes. Indeed, step 5 is often organized as homework; and sometimes the problem turns out to be trickier than expected before and cannot be solved in about 20 or 30 minutes. One of the obstacles the interpretation of the videos has to face is that there is virtually no information on the context of the respective classroom. The fact that we have to deal with the document as it is may lead to misinterpretations of the lessons themselves, not to speak of the respective educational systems.

When we look into our classrooms, we see at first glance that the Japanese and the German lessons follow the pattern at large and in detail:

1. The 'videos' subheadings like "Linking yesterday's lesson topic to today's topic" (JG) and "Reviewing (or Revising) previous material" (GA; GG) indicate a phase of recalling *what we know already*. And even "Sharing homework" (passim) serves this very purpose. By juxtaposing those introductions with moves such as "What we are going to do now is..." the teachers at least allude to the gap between "knowing" and "not yet knowing". Sometimes both sides are linked explicitly: "Our study today will use this as a foundation" (JG), or non-verbally by simply writing the task on the blackboard (GA).
2. In both of the American classes there is a kind of sharing homework, too. But what I do *not* find is any indication of a *new* problem to be solved which up to then could not have been solved. No difference is established which might stimulate *classroom* work; work simply proceeds with tasks "similar to the ones they worked on the previous day" – as a comment on a video tells us. This is not to say that individual students may not have had the experience of not yet knowing. On the contrary, in the US arithmetic class we observe several students asking the teacher to help them as they are not able to solve their problem. But this is not yet *classroom* work in the sense of our definition and according to the logic of the process; it is *just work* within *the walls of a classroom*.
3. Another point to be observed is *posing the problem*, n. b. a *new* problem. The problems posed in the US classrooms are not new ones. The lessons remind me of what I may call a *loop*: When the teacher becomes aware that the application does not work sufficiently, he or she goes back and poses the problem once again. These loops can be observed during the phase of "Sharing Homework" where we see that yesterday's problem has not been solved yet; it is still a "new" problem for some of the students.
4. Regarding the logic of the classroom process, it does not matter at all what kind of problem is to be worked on, this differs from the focus of the TIMSS studies which is on the students' intellectual level. The TIMSS Video studies have classrooms in their focus and not individual students.
5. *Tools* are provided: "What we know already", serves as a tool for further work, and so do the students' "minds" – one of the Japanese teachers puts it explicitly: "I will have everyone use their heads and think a little" (JA).

6. Work is done and at the end the *result* is stated – on the board or on a transparency.
7. The newly acquired ability is applied to similar problems: seatwork or homework in all of the classrooms.

Thus we can identify more or less explicit indications of the logic in the data representing the real processes. We can state whether or not a certain lesson complies with that general pattern – Stigler’s script. Perhaps we may speak of “good” lessons or of “bad” ones on that basis. But one ought to be careful: Any comparison or evaluation of our classrooms would not be fair as

- we virtually do not know anything about the respective contexts, even worse:
- the videos do not show the lessons in full length, and even a native speaker can hardly understand the teachers’ words in case of the TIMSS recordings. We cannot estimate the reliability of the subtitles and we cannot reconstruct the criteria according to which the videos were edited despite of what Stigler tells us in the introductory interview.

I’m afraid I must disagree strongly with Stigler when he says in his introduction that the lessons are “representative of teaching in the three countries”. The same goes for all his followers who simply jump to similar conclusions.

I doubt that it was a good idea to compose videos like those and to distribute them as an argument in political discourse or as an evidence to test far reaching hypotheses in educational research. What could be done at best is an interpretation and criticism of the individual classes’ hidden ideologies. By no means can the recordings support a comparison and evaluation of classrooms on an international scale or even within the limits of one nation.

To be fair I must add that Stigler and his team did not go as far as to link his “scripts” with the TIMSS scores directly. They were “[...] seeking to describe the classes from both the perspective of teaching practices and that of the opportunities and experiences provided for students”⁴ (Stigler et al., 1999, p. 1).

⁴ My emphasis, P. M.

So, what shall we do with such nice data? What may they be good for?

1. The model of a classroom process I developed has a *normative aspect*: Whenever a given classroom process does not match that model, it is reasonable to *question* whether the situation really is a *classroom process* and not, for instance, a mere pattern drill with restricted pedagogical concern at best. For example one could ask colleagues from the United States if the processes seen in both of the videotaped lessons really correspond with a classroom process in the sense of our definition. Isn't what we observe in the arithmetic lesson (UA) just *individual* work of students taking place in one room and supervised by one teacher who addresses all students from time to time? And isn't it just individual solving of tasks during the geometry lesson (UG) with the teacher doing nothing but determining the pace of work? So the data might allow to question single lessons critically whether or not the situation observed is really a *pedagogical* one or other.
2. The data is a treasury of *instructional options* at every point of the process – be it the ways of opening the lesson, of organizing interest in the problem; be it the complexity of the problem itself; be it the measure of the steps in the process: step by step (as in the German lessons) or a long stride (as in the Japanese ones); or be it the ways of testing the outcomes. As we see nowadays there is detailed and sophisticated research into such problems in consequence of the TIMSS video study and its follow-up studies in the US and abroad:
 - studies on proof and argumentation;
 - a multidisciplinary analysis of the negotiation of meaning;
 - the identification and scrutiny of “teachable moments” and their components; and the like.
3. The tapes as well as the findings of that research may be considered a *quarry of didactical tools*. Thus

the TIMSS 1999 Video Study of eighth-grade mathematics lessons [that] begins where the 1995 study ended ... [was] based on the premise that the more educators and researchers can learn about teaching as it is actually practiced, the more effectively educators can identify factors that might enhance student learning opportunities and, by extension, student achievement. (Gonzalez et al., 2004)

But let us be careful: our videos teach us that the tools and their composition in a real classroom are not culture-free. Can we, for example, really understand the willingness to work or the complexity of the problems worked on in the Japanese classroom without considering the teacher's role and the meaning of "learning" in Japanese society? Or one cannot understand the German lessons without knowing something about the tradition of reform pedagogy and of the Socratic method in the German *Didaktik tradition*. Nevertheless, the way the Japanese teacher makes sure that all of the students understand the problem he posed might provide a source of inspiration of German or US teachers as well.

Conclusion

Generally speaking, we have to interpret the data in its variety *within the framework* of the respective societies – their ruling beliefs, the roles of professionals and the impact of education outside school.

The first point – and I think it is the crucial point – is *teacher education*. What kind and what amount of knowledge of education, curriculum, and classroom is conveyed in what amount of time and in what kind of institutions in the respective countries? Furthermore, the case of Japan teaches us that we have to take not only the pre-service but also the in-service training into account.

The next point is the *role of a teacher* in school and in the society. The German arithmetic teacher would apparently like to be the students' older friend. At the same time he makes use of numerous disciplinary measures in order to ensure the course of the classroom process for which he feels and actually is responsible. On the other hand, the Japanese teacher is able – within a firmly established framework that is symbolized by the welcoming ritual – to play a clearly defined and obviously accepted part as a teacher in the narrowest sense of the word.

The role of the teacher is just one aspect. As is well known, learning outcomes are heavily influenced by the *individual teacher*: something like a German teacher does not exist just as well as, for example, a Czech teacher as such.

Furthermore, the *status of mathematics* – subject matter in general – has to be taken into account. Is math closely connected to related school subjects or is it isolated? In which way does school organization make allowances for the fact that there are differently gifted students? Is there a sort of *external*

differentiation according to the levels of the students' competences such as we find in Germany, or a commercialized system of private lessons like the Japanese *juku*?

Finally, it is not only school. What is the significance of the *family* for school and teaching? What precepts of the value of knowledge for their lives in society do the students have in mind when they start school? Do the views of the family on the one hand and those of the school system on the other concur or compete with each other? Moreover, we know that affiliation with a certain *social class* is a significant determinant of success in school.

My comments are anything but explanations, and the points that I have referred to are anything but variables in the technical sense of the word. What I want to stress is that the data produced in the original *TIMSS videotape classroom study* as well as in its followers do not support the proof of the TIMSS hypothesis, namely: *Different and particularly nation-specific classroom cultures may cause differences in average test scores.*

Apparently the authors of the main study came to the same conclusion. According to Gonzalez et al. (2004), they state on the basis of the studies' findings:

One of the questions that prompted the 1999 study was whether countries with high achievement on international mathematics assessments such as TIMSS share a common method of teaching. But the results from the 1999 study of eighth-grade mathematics teaching among seven countries revealed that, among the relatively high-achieving countries, a variety of methods were employed rather than a single, shared approach to the teaching of mathematics. (Gonzalez et al., 2004, p. 2)

The *cultures* or national *scripts* may influence test scores – among other “variables” that are not and cannot be controlled by video recordings. So, once more, what might the videos be good for? In short, during their training, teachers and particularly teacher students have to be provided with the *experience* of teaching and classroom management – that is what *videos* of different classrooms may convey. And they have to elaborate a notion of *classroom* that may guide them when it comes to understanding what they can see and experience – that is what *working on* the videos has to aim at. I fully agree with Stigler here who recommends us to study teaching abroad which might help us to see our own teaching more clearly, to understand, discuss, and to begin to be able to talk about teaching, in other words, talking about *what teachers actually do* in the classroom.

Acknowledgement

I would like to thank Olaf Schauerte who helped me to translate the first version of this paper into English.

References

- Stigler, J. W., Gonzalez, P., Kawanaka, T., Knoll, S., & Serrano, A. (1999). *The TIMSS Videotape Classroom Study: Methods and Findings from an Exploratory Research Project on Eighth-Grade Mathematics Instruction in Germany, Japan, and the United States*. Washington: U.S. Department of Education. Retrieved from <http://nces.ed.gov/pubs99/1999074.pdf>
- Gonzalez, P., Guzman, J. C., Partelow, L., Pahlke, E., Jocelyn, L., Kastberg, D., & Williams, D. (2003). *Highlights From the TIMSS 1999 Video Study of Eighth-Grade Mathematics Teaching*. Washington: U.S. Department of Education. Retrieved from <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2005005>.

Author

Prof. Dr. Peter Menck, University of Siegen, Faculty of Arts, Adolf-Reichwein-Str. 2, AR-K 227, 57076 Siegen, Germany, e-mail: menck@erz-wiss.uni-siegen.de