TALIS Video Study

Introduction

In this paper we will outline some basic elements to be considered when planning a video design for investigating instructional practices among OECD countries. The paper, and upcoming report, builds on the extensive amount of video research conducted over the last years (see table 1), and link this research to recent methodological and technological developments in observational studies of teaching and instruction.

Video studies have proven to be a valuable tool for investigating instruction. By aggregating primary data consisting of video recording of classroom teaching, video analysis allows for identification of country –specific patterns of instruction, so- called cultural scripts (see Stiegler & Hiebert, 1999; Roth et al., 2006). It also enables identifying cause – effect relations in different teaching- learning scenarios (Seidel & Prenzel, 2005; Fischer & Neumann; 2012,) and gives opportunities for in- depth analyses of instructional processes (Clarke et al., 2006; Klette, 2009; Borko et al., 2008).

Although the TIMSS 1999 Video Study illustrated some of the benefits of video studies, the costs of participation severely limited its reach and impact. Much has however changed since then, and some of the most important new developments are summarized in these dot points:

- New technologies allow for easy storage and online streaming of video.
- Video recording equipment is miniaturized and made portable, and can be remotely controlled. It can also be operated by individual teachers, and do not require a cameraperson
- Developments in tools for sharing observations and collaborating on analyses such as common coding manuals and new software tools targeted towards analyses of video data facilitate comparative and collaborative analyses.
- Many video studies have been carried out during the last decade, both small scale studies of individual teachers and more comprehensive large scale studies. Particularly the latter are building on increasingly refined models of instructional quality. Establishing a closer connection between these kind of studies have the potential of greatly increasing our knowledge of key criteria for instructional quality.
- Developments in nested research designs, like Hierarchical Linear Models (HLM) or Structural Equation Models (SEM), make it possible to analyze the data on different levels, i.e. individual level, school level, national level and to investigate correlations between instruction and student achievement.

Based on these points, we believe that a new wave of video studies has extraordinary potential. We will argue that a theory- based video analysis design, integrating quantitative and qualitative measures within a coherent framework, have the potential of providing data that through detailed analysis will largely expand our knowledge of key criteria of instructional quality. Furthermore, this should give us new insights into the relationship between high quality instruction and student achievement. In order to achieve such goals, we should re-conceptualize our study designs, taking advantage of both recent technological developments and developments of analytical instruments within the field. By using remotely operated and portable miniaturized recording equipment for data collection, common coding manuals for analysis, and advanced software tools for representing and analyzing video data, we can analyze much richer data sets than previous studies, and to a much lower cost. If participation in the project is no longer cost prohibitive, more countries can be included, thus amplify both the richness of the data and the extent of its impact. If we leverage new technologies effectively it will be possible to create less centralized, more collaborative research process, and to provide more constituencies with data access in more varied ways. In addition, nested research designs provide opportunities for analyzing both different levels of data and to integrate different kinds of data. Through the use of HLM models and EMS models it is possible to investigate correlations between quality of teacher instruction and student performance. In summary, we will argue that a theory- based video analysis design, integrating quantitative and qualitative approaches and using common coding manuals to measure qualities of instruction, makes it possible to systematically explore instructional qualities across national contexts and settings. This is a huge step forward compared with many other previously conducted research studies. In studies using predominantly self-report instruments (TALIS; TIMSS; PISA), providing little or much lower quality instructional data, no direct relation can be established between teaching and learning processes.

Review of relevant video studies

Video recording has been acknowledged as a promising way to decompose qualities in teaching and instruction, and in different countries in Europe, Australia and the US video research has been a common way to investigate classroom learning and teaching. In this short overview we will give a brief description of some significant video studies relevant for the TALIS Video Study. The TIMSS Video Study stands out as a pioneer study in this respect, but also the IPN Video Study (Germany), the CPV Video Study (Czech Republic), the LPS Study (Australia), the MET study (US), and small scale studies like the Problem Solving Cycle/STAAR study (US), the PISA + Video Study (Norway), and the GestePro study (France) represent agreeable ways of conducting and exploring studies of classroom processes.

Some of the above mentioned studies might be described as explorative and inductive (LPS study) while others are theory driven (MET study) aiming to provide a composite indicator of effective teaching. The studies have a broad variation in study design, research ambitions, focus of research, school subjects included, sample size, number of video filmed lessons etc., (see table 1).

As can be seen in table 1, some of the studies focus on only on one school subject, while others analyses and compare across school subjects. Furthermore, while some of the studies focus on instructional practices exclusively (i.e. TIMSS study; LPS study), others also use video documentation for professional development purposes (STAAR study, IPN study:

Observe). The studies further vary in terms of how clearly they have developed an overall framework for analyses and common coding procedures. While the LPS study might be described as a comparative design for data gathering procedures, the PISA + Study, the IPN Study as well as the TIMSS Video Studies developed common procedures for data analyses. The ongoing MET Study in addition use common coding manuals as instruments to test out possible tools for teacher evaluation.

The studies further vary in sample size (small scale video studies and large scale studies) as well as the amount of lessons videotaped in each classroom. While the TIMSS and the IPN study have developed criteria for a representative national sample, covering one lesson (or a double lesson unit) in each classroom, the LPS study used "competent teachers" as agreed upon by local criteria, as selection criteria and video filmed 10 consecutive mathematic lessons in all classrooms. Last but not least the studies vary in how they are linked to outcome measures. Table 1 summarizes some of the key elements that vary across the studies.

Table 1: Key elements that vary across recently conducted classroom video studies

Title	**Research focus**	**Subjects in focus**	**Countries involved**	**Sampling procedures**	**Sample size (classrooms pr country)**	**Lessons video recorded in each subject**	**Coding procedures**	**Outcome measure**
TIMSS 1997 (US)	Instructional practices	Mathematics	3	Probability sample	70- 100	1 lesson	Common coding manuals	None
TIMSS 1999 (US)	Instructional practices	Mathematic Science	7 5	Probability sample	70 - 100	1 lesson	Common coding manuals	None
IPN study (Germany)	Instructional practices	Science (Physics)	1	Probability sample	50	2 lesson unit	Common coding manuals	Pretest/ posttest
CPV study (Czech Rep)	Instructional practices/PD	Physics, Geography, English, Physical Ed.	1	Non-probability sample	30 25 40 27	2 lesson unit	Common coding manuals	None
Swiss/German study	Instructional practices	Mathematics	2	Probability sample	40	Consecutive Lessons (3)	Common coding manuals	Pretest/ posttest
LPS study (Australia)	Teaching-learning interaction	Mathematics	14	Non-probability sample	3	Consecutive Lessons (10)	Selective coding	Test*/ TIMSS items
PISA + Study (Norway)	Teaching-learning interaction	Mathematics, Science & Reading	1	Non-probability sample	6	Consecutive Lessons (7-9)	Common coding manuals	None
GestePro (Tiberghien) (France)	Teaching-learning interaction	Science	1	Non-probability sample	7	6 lessons during one month	Inductive coding	None
QuIP (Germany/ Finland/Swis)	Instructional practices	Science (Physics)	3	Probability sample	50 -Ger 30 -Swiss 30- Fin	Consecutive Lessons (3)	Common coding manuals	Pretest/ posttest
STAAR study (US)	Instructional Practices/PD	Mathematics	1	Non-probability sample	8	Lesson units Throughout the year	Common coding manuals	None
MET study (US)	Instructional practices	Mathematics, Science & Language art	1	Probability sample	3300	4 non-consecutive lessons	Common coding manuals	Several test scores

* Test scores not used in the final analyses in the LPS study

It should be noted that all the above mentioned studies also collected additional data for analytical purposes through for instance: teacher questionnaires, student perception surveys, video stimulated teacher and student interviews, cognitive assessments, field notes, collection of learning artifacts used during the lesson, copying of student work, assessment of teachers pedagogical and/ or content knowledge. These are not included in table 1.

In designing a TALIS Video Study, it seems particularly important to closely study all aspects of previous large scale video studies like the ones here presented. However, the maybe most important element for making decisions related to study design, is to clearly define the research ambitions of the TALIS Video study.

Research ambitions of the TALIS Video Study

Two major options for describing the research ambitions of the TALIS Video Study could be described as follows:

i) National representative research ambitions – that is, to design a study that makes it possible to identify and compare instructional practices across or between nations and link them to student proficiency levels.
ii) In depth studies of illustrative classrooms practices from each country – that is, to only make comparisons between and across classrooms, with no ambitions of making comparisons across or between nations.

The first option will provide data that through analysis will have the potential of giving us an overview of instructional practices in mathematics classrooms at this level and link these practices to student proficiency levels. The TALIS Study will then report representative case descriptions from the participating countries, and provide educational policy makers with crucial information in this prioritized key area. A decision of only pursuing the second option will limit the TALIS Study to primarily report illustrative contextualized case descriptions of mathematics classrooms from the participation countries.

If, however, research goals are limited mainly to document differences in instructional practices between and within countries, and with modest ambitions of linking these practices to student outcome data, probability samples are not needed.

Protocols for sampling and data collections in the TALIS Video Study

Given a decision of choosing the most ambitious research goal (i), probability samples should be collected. This could be done in various ways, and a detailed description of this question will be outlined in the upcoming report. However, a few dilemmas related to decisions about sampling procedures will be presented. Probability samples necessitate a relatively high number of video filmed classrooms in each country (i.e.25-30), but relatively few lessons in each classroom would be needed (3-4). As an example, 3 lessons from 30 different classrooms in each participating country, 90 lessons of mathematics instruction in each country, will be sufficient for applying HLM modeling of analyses. This should be feasible in relation to costs and time schedule. Reducing the number of participating classrooms in each country will require a larger number of videotaped lessons from each of the classrooms (for example 5 lessons). Assuming 12 classrooms from each country, this equals 60 lessons of mathematics instruction from each country. Even if this second solution probably will reduce the total costs somewhat, it impedes the possibility of performing representative national analyses thereby reducing the potential significance of the analyses considerably.

It should also be taken into consideration that the most cost-demanding part of the TALIS Video Study will be to carry out the data analyses required for making cross national comparison, even if these expenses to a certain degree will vary according to national wages for this kind of work.

Student achievement measures

Varies options for measuring student achievement in the TALIS Video Study seems possible. Using pre-tests and post-tests, as for instance was done in the IPN-Video Study, is one alternative. Developing such tests is quite costly, but the expertise needed would be available within research groups already working with other OECD studies (PISA). A different solution could be to make an even closer connection to PISA. One could for instance make tests using items from previous PISA studies (2003 - 2012), or coordinate the TALIS Video Study with the PISA 2015 Study, that is, give the students in the selected video filmed classrooms the PISA 2015 test. A carefully designed representative sample of 30 classrooms from each of the participating countries also makes it possible to use the national PISA score in mathematics as a measure for student proficiency and to discuss instructional practices in relation to national PISA scores. Generally, using PISA test items should be cost effective and in addition have some obvious advantages related to data analysis, as benchmark levels and scales are already established.

Costs to be considered

A budget for a TALIS Video Study must therefore take into account the following three elements:

1) Costs linked to the data gathering process
    - Video recording equipment (i.e. portable boxes a 8 000,00 EURO)
    - Videotaping lessons in classrooms (i.e. remotely controlled solutions)
2) Costs linked to developing student achievement tests (alternative 1: using existing measurement scores like the PISA score/ alternative 2: to develop own pre/ posttests)
3) Cost linked to analyzing the data (common coding manuals/ training facilities etc.)

A more detailed description of main aspects of budget and financing will be developed in the full report.

Thematic focus – mathematics instruction

We assume that the idea of conducting a TALIS Video Study originates from the wish to study instruction more closely and to relate the findings to other educational OECD studies, i.e. TALIS and/ or PISA. A TALIS Video Study should make it possible to not only observe what is happening in the classroom in relation to the themes covered (What), but aim to identify the actions, procedures and behaviors that support student learning in this area (How). Based on this we suggest that mathematics instruction should be the focus for a TALIS Video Study. Besides being an area of uttermost importance for skills for the 21st century, mathematics instruction is an area with interesting patterns of variation across the OECD countries. Mathematic instruction is also relatively stable in terms of thematic areas, content coverage, and national curricula which make comparative analyses possible. Furthermore several studies (Hill et al., 2005; MET, 2010) suggest that teachers seem to have a larger influence on students' math performance than other subjects. To identify high quality mathematics teaching across national contexts would therefore be especially important. Last but not least, a TALIS Video Study on mathematics would make comparisons with the TIMSS 1999 and 2003 video studies possible (reuse of items for analyses).

A relevant mathematical theme could be the teaching and learning of statistics. Depending on the final decisions about the number of lessons to be video filmed in each classroom, this theme could be narrowed down further.

Developing a common analytical framework

Video data is a kind of raw data and allows for analyses from multiple perspectives and theoretical positions (Klette, 2010; Tiberghien & Sensevy, 2012; Fischer & Neumann, 2012), and for comparative analyses a common theoretical framework and common coding manuals is required. A common coding manual serve two functions

- Makes it possible to compare characteristics of teaching learning processes systematically
- Contribute to develop reliable measures for analyzing instructional practices

The LPS study developed a sophisticated research design for gathering data within a comparative framework which made it possible to compare dimensions of math education across settings and national settings (interaction patterns, cognitive activation, the role of peer learning, quality of the tasks etc.). However due to a lack of common coding manuals – the whole LPS material were never analyzed in terms of instructional practices across all participating countries. One of the strength of the ongoing MET study (Measures of Effective Teaching) (http://www.gatesfoundation.org/college-ready-education/Documents/preliminary-findings-research-paper.pdf) is how this study systematically evaluate and test out the effects of different coding manuals, that is both generic and content specific coding manuals.

We will argue that the upcoming TALIS Video Study should use common procedures for both data gathering and data analyses. Analyses based on the common coding manuals might represent a minimum version of the required analyses from the participating countries. In addition video data, supported with assignment collections and/ or copies of students work from each of the videotaped classrooms allow for multiple additional analyses such as in depth analyses of individual teacher profiles, the role of peers, content activation and cognitive demands linked to the different tasks – all these dimensions scholars agree to have crucial impact for quality instruction in mathematics.

For analyses, a valid and reliable coding procedure has to be decided on and supported with training facilities. During the last couple of years considerable work has been done in trying to test out the possible differential strength and weaknesses linked to the different coding manuals (Grossman et al. (2010). The MET study for example uses existing coding manuals (MIQ (mathematics); PLATO (language art); Charlotte Danielson framework (generic); CLASS (classroom climate) to systematically test out the relative strengths and weaknesses linked to the different manuals. Instead of developing new coding manuals, which is a complex and time consuming process, we suggest that the TALIS Video Study modifies existing coding manuals to test out possible differential instructional practices across the participating countries.

References:

Clarke, D., Emanuelsson,,J., Jabblonka E., & Mok, (2006) *Making Connections: Comparing Mathematics Classroom arpound the world.* Rotterdam: Sense Publiserhers.

Borko. H., Jacobs,J, Eiteljorg, E, & Ptitmman,M.E. (2008). Video as a tool for fostering productive discourse in mathematics professional development. *Teaching and Teacher Education, 24,* (pp 417-436).

Hill H., Rowan, B.D., Ball D.L. (2005).Effects of Teachers' Mathematics Knowledge for Teaching and Student Achievement. *American Educational Research Journal, Vol. 23* (pp371-406)

Hill, H.C., Blunk, M., Charalambous, C., Lewis, J., Phelps, G.C., Sleep, L., & Ball, D.L. (2008). Mathematical Knowledge for Teaching and the Mathematical Quality of Instruction: An exploratory study. *Cognition and Instruction*, *26*, 430-511.

Fischer H., & Neumann, K. (2012). Video Analyses as a Tool for Understanding Science Instruction. In Jorde & Dillon (Eds.). *Science Education Research and Practice in Europe*.Sense Publishing (pp 115- 138)**.**

Klette, K. (2009). Challenges in Strategies for Complexity Reduction in Video Studies. Experiences form the PISA + Video Study: A Video Study of Teaching and Learning in Norway. In T.Janik & T. Seidel (Eds.) *The Power of Video Studies in Investigating Teaching and Learning in the Classroom.* Munchen: Waxmann.

MET (2010): Learning about Teaching – preleminary findings (http://www.gatesfoundation.org/college-ready-education/Documents/preliminary-findings-research-paper.pdf)

Roth, K.J., Druker,S.L., Garnier,H.E., Lemmens,M., Chen,C., Kawanaka,T., Rasmussen,D., Trubacova,S., Warvid, D., Okamoto,Y., Gonzales,P., Stiegler,J., & Gallimore,R. (2006). *Teaching science in five countries: Results from the TIMSS 1999 Video Study* (NCES 2006-011), Department of Education, National Center for Education Statistics, Washingotn DC: US Government Printing Office.

Stiegler J.W., Hiebert,J. (1999).The teaching gap: Best ideas from the world's teachers for improving education in the classrrom. New York: Free Press.

Seidel T., Prenzel M.(2005). *How to run a video study. Technical Report of the IPN Video Study.* (Kobarg M, Ed). Munster:Waxmann.