



Third International Mathematics and Science Study

**TIMSS 1999 VIDEO STUDY**

**TRANSCRIPTION/  
TRANSLATION  
MANUAL**

## TABLE OF CONTENTS

<i>Overview</i>		3
	TIMSS 1999 Video Study	
	Goal of the Study	
	Importance of Standardized Transcription Procedure	
<i>Chapter One: Transcription/Translation Procedure</i>		5
1.1	Identifying the Speaker	5
	Speaker Codes	
	Blackboard	
1.2	Transcription Conventions	6
	Change in Speakers	
	Overlapping Speech	
	Punctuation, Diacritical Marks, and Other Conventions	
1.3	Common Backchannels	11
<i>Chapter Two: Second-Pass Transcription Procedures</i>		13
2.1	Items You Will Need to Transcribe/Translate/ Timecode the Videodata	13
2.2	Step-By-Step Procedures	13
	Step 1: Check lesson assignment schedule	
	Step 2: Access file	
	Step 3: Review Transcription/Translation	
2.3	Transcript 1 and Transcript 2 Text Fields	14
2.4	Problems	14
2.5	Editing	14
	Three-Line Segments	
2.6	Timecoding	15
2.7	Logging Hours	15
2.8	Completing a Unit	15
<i>Chapter Three: Transcription/Translation Issues</i>		16
3.1	T/T Meetings	16

## Overview

### TIMSS 1999 Video Study

This manual describes the protocol used for transcribing/translating the video data collected for the Third International Mathematics and Science Study, known as TIMSS 1999 Video Study.

TIMSS 1999 Video Study is a cross-national study of eighth-grade mathematics and science classrooms. The study involves videotaping and analyzing teaching practices in more than one thousand classrooms in six countries.

The study is funded by the U.S. Department of Education and the National Center for Education Statistics (NCES) and conducted by LessonLab, Inc. of Los Angeles, California.

### Goals of the Study

The TIMSS 1999 Video Study has the following six goals:

- ◇ To investigate mathematics and science teaching practices in U.S. classrooms.
- ◇ To compare U.S. teaching practices with those found in high-achieving countries.
- ◇ To discover new ideas about teaching mathematics and science.
- ◇ To develop new teaching research methods and tools for teacher professional development.
- ◇ To create a digital library of images to inform U.S. educational policy.
- ◇ To stimulate and focus discussion of teaching practices among educators, policy makers, and the public.

To achieve our goals we will collect data from 100 randomly sampled schools in each participating country. There are three sources of data: videotaped lessons, teacher and student questionnaires, and supplementary materials (e.g. worksheets, copies of textbook pages) used by the teacher and students during the videotaped lessons.

Data collection. From each of the 100 schools, we will videotape one mathematics and one science lesson. All are eighth-grade lessons. Each lesson will be videotaped using two digital camcorders by one videographer, one focusing mainly on the teacher and the other on the students.

Data processing. Both videotapes will be digitized and burned onto CD-ROM. The tapes that focus on the teacher will be copied to VHS tapes and sent to transcribers/translators. The lessons will be transcribed from the original language into English, and sent back to LessonLab as Microsoft Word files. The transcripts will then be timecoded, linked to the video footage, and stored in a multimedia database.

Data analysis. The video data will be reviewed by our research group who will code various aspects of lessons to describe mathematics and science teaching in each country,

and the results of the coding will be analyzed statistically and reported to NCES by LessonLab.

### **The Importance of Standardized Transcription Procedure**

The transcript is not intended to be a replacement for the viewing of the actual videotaped interaction; however, the transcript will serve as an aid in the coding and subsequent analysis of classroom interactions. Therefore, the transcription of the videotaped data should reflect, as accurately as possible, the words spoken by both teachers and students alike. It is imperative that this process be executed in a careful and consistent manner, as it is crucial to the success of the entire project. Therefore:

Do not summarize or paraphrase what the participants are saying.

Do not translate/transcribe the data to interpret what the speaker "meant to say."

## Chapter One: Transcription/Translation Procedure

In dealing with face-to-face interaction between teachers and students, various issues will arise that may, at first, seem insignificant. However, phenomena such as overlapping talk and pauses will have an impact on how the data will be coded and analyzed.

Though every classroom interaction will have its own unique qualities and attributes, there is sufficient similarity between these interactions to create a consistent system for the representation of the interactions in a transcript. This manual details this system and explains all of the transcription conventions you will be using in your work.

### 1.1 Identifying the Speaker

In a normal situation, there is one teacher and a group of students in a classroom who are doing most of the talking. The teacher's voice will be captured by a radio microphone that will be wired to the teacher. The microphone sends a mono-signal to the first camera. This radio mike will also capture the students' voices who are a few feet away from the teacher. A stereo microphone mounted on the same camera will capture the voices of students who are farther away from the teacher. This stereo microphone sends another mono-signal to the camera. When you listen to the dialogue with a headphone, you will hear the audio from the teacher's radio mike on one side and the camera-mounted mike on the other side of the headphone. When the audio from the two sources is confusing, you may need to listen to one side of the headphones at a time.

#### *Speaker Codes*

The following five speaker codes have been developed to try to deal with, as reasonably as possible, the discourse of the classroom:

<b>T</b>	Teacher
<b>S</b>	Single student
<b>SN</b>	Student-new. A single student whose identity differs from the last student to speak
<b>S?</b>	When the identity of the student (whether the speaker is S or SN) is unclear
<b>Ss</b>	Multiple students, but not the entire class
<b>E</b>	Entire class (or sounds like the entire class); used to indicate choral responses
<b>O</b>	Other; used to indicate speech by a non-member of the class, such as school personnel, office monitors, or talk from public address systems

While it is generally easy to distinguish the teacher's voice from that of the students', it is not always possible to distinguish between individual student voices. Thus, there will be no attempt to track the voice of any individual student in the ongoing discourse (e.g., marking Student 1, Student 2, Student 3, etc.).

- Whenever *more than one student* (but not all) speak simultaneously, the code **Ss** is entered.
- In order to keep track of a single student's continued speech, an **S** should be used to indicate a given student's speech until another student begins to speak. The next student should be indicated with **SN** (Student-new). If the new student continues to speak, an **S** should be used to mark their next turn. **SN** is always used to indicate that a new student (i.e., a student whose identity differs from the previous student speaker) is speaking. The very first student to speak in a lesson is identified with the **SN** code.
- Whenever it sounds as if *the entire class* speaks simultaneously (i.e., a "choral" response), the code **E** is entered.

It is important at this point to understand what is meant by "simultaneously" in this particular setting. Simultaneous talk refers to those occasions when more than one student is speaking (1) *the exact same word(s)* (2) *at the exact same time*. Both of these conditions must be met in order to use either the Ss or the E speaker code.

### ***Blackboard***

There is one additional code used to mark material written on the blackboard:

**B** Blackboard; used to indicate the translation of foreign words/phrases written on the blackboard

The code **B** is used to indicate the translation of foreign words/phrases written on the blackboard. Only those **words** or **phrases** that cannot otherwise be recovered from the ongoing discourse are to be translated and added to the transcript.

For example, if a teacher writes a word problem on the blackboard but does not say the word problem out loud or otherwise indicate what the problem is about, then this should be translated and added as a separate turn in the transcript.

If, however, a teacher writes a word problem (or other phrase) on the board and then says it out loud to the class, the translation of the written word(s) should not be added to the transcript as it is fully understandable by following the transcript of the teacher's discourse.

Formulae written on the blackboard which contain only numbers or other mathematical notations (e.g.,  $x$ ,  $y$ ,  $\beta$ ,  $\_$ ) should be added as a separate entry into the transcript (e.g.,  $2x+3=10$  does not need to be "translated") ONLY the first time a problem is discussed. After the initial reference to a problem, it is not necessary to transcribe it again. Numerical problems that are not discussed do not need to be transcribed.

## **1.2 Transcription Conventions**

Take a look at the sample transcript below. The letters to the left of the text are the speaker codes. Notice in the text various symbols are used.

T So by looking at- this one is ... eight square units this one is twelve square units this one's twenty-four square units how do I find the area of

a rectangle? Now find me a mathematical way of doing it. So I don't have to count all the time. BRIAN?

S Multiply the vertical squares times the horizontal squares.

T That's exactly correct ... Who could state it in another way?

T Anyone //else?

S //Well maybe...

Besides using the same speaker codes to transcribe the videodata, it is necessary that all transcribers use transcription conventions in a standardized manner so that they convey the same meanings. This section explains all the conventions that are to be used in transcribing the videodata for our project.

### ***Overlapping Speech***

Each speaker contribution is to be entered as a separate turn at talk if two speakers are speaking at the same time, but they are saying two different utterances.

Likewise, if two speakers are saying the exact same thing, but are not speaking at the exact same time (e.g., one student begins to speak in the middle of the other student's utterance), this does not constitute "simultaneous" speech. Although the two speakers are saying the exact same words, such moments will be treated as cases of "overlap," and each speaker contribution is to be entered as a separate turn at talk.

In the transcription system adopted here, moments of overlapping speech will be indicated by double backslashes (//) to indicate where the overlap begins, as in the following example:

T All right giggle boxes. What is so silly?

S When you //were holding up a ten Jordan was holding up a ten.

SN //When you were holding up a ten Jordan was holding up a ten.

The double backslashes should always come in a set. The point where the two speakers begin to talk over each other should always be marked in both the overlapped turn (the one who began speaking before being overlapped, as in the second line above) and in the overlapping turn (the one who began speaking during the talk of another, as in the third line above).

#### One Exception to the Simultaneous Speech Code

Since the speaker code **Ss** does not include the teacher, the teacher's talk should always be treated as a new and separate turn, regardless of whether the teacher overlaps or speaks simultaneously with another interactant.

One important exception to the "simultaneous speech" rule concerns when the teacher and a student (or students) speak the same words at the same time (a very rare occurrence). In such a case, the teacher's utterance should ALWAYS be entered as a

separate turn from the student's turn. Double backslash (//) should be used to show that the teacher and student overlapped in speech.

For instance, in the following example, although the teacher and the student are saying the exact same words at the exact same time, their respective turns at talk should not be combined as one turn, but as two separate turns:

T      What's five plus one?

S      Seven.

T      //No.

S      //No.

### ***Entering a New Speaker Code: Lengths of Pauses and Activity Shifts***

**Pauses.** If a pause lasts longer than three seconds, a new speaker code should be entered, even when the same speaker resumes speaking. Organizing the speech in this manner will allow timecodes to correspond with the actual utterances in the video.

In the following utterance, a new speaker code should be entered after the pause:

T: If you turn to page five, you will see...[4 seconds]

T: This page has several problem sets that relate to your homework.

**Activity Shifts.** Several activities occur in the classroom, for example: the teacher's lecture, student group work, question and answer, students at the blackboard, etc. If such a shift occurs during an utterance, a new speaker code should be entered.

T: So that's how you would solve such a problem.

T: Now, please take out your worksheets and get into your groups.

### ***Punctuation, Diacritical Marks, and Other Conventions***

The use of punctuation marks (such as a comma, period, colon, etc.) in the transcription of videotapes in this study will follow the normal rules used in written English. In addition the following conventions will be used as part of the transcribing system:

<p><b>[Bell.]</b> Please indicate when the bell rings at the very beginning of the and at the end of each lesson by putting the word "bell" in brackets.</p>
--



**PROPER NAMES** Proper names of teachers, students, and school personnel must be changed due to confidentiality issues.

The transcripts currently use real names, which are in all CAPS:

T Okay, for the perimeter you add up the sides. How many sides are there?  
JESSICA?

Change the name using the same first letter, if possible, and delete all caps:

T Okay, for the perimeter you add up the sides. How many sides are there?  
Jane?

**- (HYPHEN)** A hyphen indicates that a speaker has "cut-off" (or self-interrupted) his/her speech.

T So, I- if I wanted- if you wanted to give me nine of those

**? (QUESTION MARK)** A question mark indicates that the utterance is to be understood as a question (usually determined through intonation), as in the following.

T Why can't the five unifix cubes be stuck together?

S Uh cause they're one.

T What?

**. (PERIOD)** A period marks the end of a phrase, a sentence, or a turn at talk that is NOT to be understood as being a question.

T Ah ha. Okay. I'd like you to put all of your unifix cubes on your name tag. So your place value chart should be empty.

When a period is placed in the middle of a turn, insert two blank spaces after the period.

**OTHER PUNCTUATION MARKS** Other punctuation marks such as commas, exclamation points, semi-colons, and colons may be used when appropriate.

T Be quiet!

**... (THREE DOTS)** A series of three dots, separated by a blank space before and after, is used to indicate a pause.

T So if you had ... is this your hand?

**( )** Empty parentheses indicates that some speaker has spoken, but the words cannot be made out.

T When you measure these two points, can you then figure out the direction?

S Yes. ( ).

**(a word or words)** Word(s) surrounded by parentheses indicate that the transcriber has made a best guess at what the speaker has said, but cannot guarantee it.

T The fixed amount is one hundred. What are they paying a day?

S Ah, (I don't know).

**(word A/word B)** If there are two alternative hearings, both of these should be included within single parentheses separated by a single backslash.

T This is seven thousand. So this line is?

S (Sixty/sixteen).

T Good.

**Numerals** Numbers should always be written out as words, in the way in which they are said (e.g. "two" instead of "2").

T How many did we start with?

S Two.

**Capital** Capital letters should only be used with proper nouns (names, cities, countries, languages, etc.), at the beginning of a new turn at talk, or after a period or question mark.

T Remember your test is on Friday? After your field trip to Palm Springs. Okay?

**(Exception:** For first pass translation, remember to put the names of the teachers and students in ALL CAPS.)

**ALL CAPS** When speakers refer to points, lines, angles, etc. by their alphabetical label, these labels should be transcribed in capital letters, even if it would otherwise appear as a lowercase letter.

T So this triangle ABC, write a little A here. Between the angle A and B.

**C-A-P-S** When speakers spell out words, each letter should be placed in capital letters, with a dash between each letter.

T How do you spell triangle?

S T-R-I-A-N-G-L-E

<p><b>[Transcriber's Note]</b> In cases where a bracketed comment is essential to the understanding of an utterance, the transcriber may enter a note, in brackets. This convention should only be used in cases that have been discussed with the translation managers, such as ironic or sarcastic phrases, or code-switching from one language to another. No non-verbal actions should be indicated in notes.</p>
---

### 1.3 Common Backchannels

Throughout their talk, the speakers make use of "backchannel" devices, "discourse markers," and "hesitation indicators" to show that they are, for example, paying attention (e.g., mm hm, uh huh), agreeing (e.g., yeah, yep), hesitating (e.g., uh, mm), showing surprise or displaying some new understanding (e.g., ah ha, oh, ah). In order to be relatively consistent throughout the transcription, these should always be transcribed as follows:

- Ah; Uh; Um, Oh
- Ah-ha; Uh-huh; Nn-hnh; Mm-hm
- Yeah; Yep
- Okay

Foreign language equivalents should be similarly transcribed in order to remain consistent.

**Cantonese** examples:

<u>Backchannels:</u>	<u>Discourse Markers:</u>	<u>Hesitation Indicators:</u>
Mm – Mm hm/Uh huh	dak/ho la – okay	uh/er – uh
Ho yeh – yeah	gam le/gam a – well	Um/ee - Um
Hai – yep	sho yee le – so	o go yi shi hai –I mean
Hai la – Ah ha	kei ming/lei ming la – y’know	

**Czech** examples:

<u>Backchannels:</u>	<u>Discourse Markers:</u>	<u>Hesitation Indicators:</u>
Mm hm	no jo – alright	Ehm - Um
Uh huh	sup – hurry	tim myslim – I mean
Jo, Tak jo - Yeah	dobre – okay, well	
Jo, Tak jo – Yep	no – well	
	tak/tak jo – okay, well	
	tak, takze - so	

**Dutch** examples:

<u>Backchannels:</u>	<u>Discourse Markers:</u>	<u>Hesitation Indicators:</u>
Mm hm	Ja maar – Yes, but	Um
Uh huh	Heh – Right	ik bedoel – I mean
Goh – wow!, oh!, gosh!	Nou – Well	
	Dus – So	

**Italian** examples:

<u>Backchannels:</u>	<u>Discourse Markers:</u>	<u>Hesitation Indicators:</u>
E poi – And then	Vero -- Right	Eh -- Uh
Be(h)	Allora, Quindi – So	Chioè – I mean
	Insomma – y’know	
	Dunque – So	

**Japanese** examples:

<u>Backchannels:</u>	<u>Discourse Markers:</u>	<u>Hesitation Indicators:</u>
M; N (㊂) – Yeah	Ja/Soreja - Well then	Ee (→) - Well; Uhh
So-so-so – Right	Ano ne - Well/Let me see...	Etto - Um
Ee (㊂) – Yes	Ne – Right/isn’t it?	
Ee (㊂) - What?	Sorede - So	
Iya – No		

**Swiss German** examples:

<u>Backchannels:</u>	<u>Discourse Markers:</u>	<u>Hesitation Indicators:</u>
Eh hn, Ae-ae – Uh huhEi, [or]	Dann -- well	ehm, eh – um, uh
Ehm, eh -- Uh	Also -- Okay	
	Oder? -- Right? Isn't it?	
	Hä – right?, okay? or huh?	

## **Chapter Two: Second-Pass Transcription Procedures**

The videodata are transcribed in two phases: first pass and second pass. Most of our first pass work is done by our subcontractors, who use VHS copies of the videos to translate the lessons from the original language into English. The lessons are then sent back to LessonLab as Microsoft Word files. In the second pass, the transcripts are linked to the video footage in a multimedia database, and are then reviewed, corrected, and timecoded by the translator/transcribers.

In this section, the second-pass transcription/translation process will be explained. This process includes keeping track of problems, glossary terms, and transcribing hours. The following instructions will assume that training has already taken place. Please read through this section carefully and thoroughly.

## 2.1 Items You Will Need to Transcribe/Translate/Timecode the Videodata

- Work station (i-Mac)
- Lesson Folder from the Lesson Files
- Forms: Checklist, Translation/Transcription Log, Glossary/Problem Form, Name Change Form

## 2.2 Step-By-Step Procedures

### *Step 1: Sign Up for Lesson in the Sign Out Book*

Rather than pre-assign lessons to particular individuals, we ask that transcribers check the Sign Out Book to find the lessons available for second-pass work. This book is kept on top of the file cabinet and contains lists separated according to country (i.e., US, the Netherlands, the Czech Republic). Transcribers should work on the next available lesson (as noted on the Lesson List), and write their names and the date next to the lesson identification number to indicate that it is being worked on.

### *Step 2: Access the file*

All second-pass transcription/translation and timecoding will be done at a computer work station using **vPrism** software. Videos of the lessons have been digitized and stored on the server as Mpeg files. Transcripts completed by first-pass reviewers are imported into the **vPrism** database and linked to the Mpeg files on the server. You can access the **vPrism** database that is stored on the **4-D** server from your work station by opening **4-D Client** from the Apple Menu. Detailed instructions about **4-D Client** and **vPrism** are provided separately.

### *Step 3: Review Transcription/Translation*

Review first-pass transcripts and make corrections following the system outlined in the first section. Additional materials have been collected along with videotapes, (i.e., copies of worksheets, textbook pages, and overhead projection materials used in the videotaped lessons), so use them to clarify or confirm the content of conversations.

## 2.3 Transcript 1 and Transcript 2 Text Fields

The **vPrism** software allows you to transcribe in two fields: Transcript 1, and Transcript 2. All material should be transcribed in Transcript 1, with the following exception: Two simultaneous conversations where the camera is on a student to student conversation but the teacher, who is speaking to a student in another part of the classroom is still audible and overlaps with the student-student conversation. In these cases, the teacher-student talk should be transcribed in Transcript 1, while the student-student conversation should be transcribed Transcript 2.

## 2.4 Problems

While working on second-pass transcription, you may encounter some technical or translation problems. Problems can include, but are not limited to, difficulty in hearing, finding English language equivalencies for translation, or translating idiomatic expressions. Whenever you encounter a problem, please make note of it on the Glossary/Problems form. Please fill out these forms *neatly* and note the screen time before continuing with your 2<sup>nd</sup> pass work. These forms will help us keep track of problematic areas and issues for discussion. This will also help you keep track of those portions of the videotape that you were unable to translate/transcribe and that need to be re-examined once you have consulted with others.

## 2.5 Editing

### Three Line Segments

As part of the 2<sup>nd</sup> pass. Long turns should be divided into segments of three lines or less. This is because completed transcripts are later displayed as subtitles on the video, and the video frame can only contain three lines of text. In the following example, the teacher's original utterance (a) is displayed as two separate turns in the transcript (b):

(a) Before editing:

T So by looking at- this one is ... eight square units this one is twelve square units this one's twenty-four square units how do I find the area of a rectangle? Now find me a mathematical way of doing it so I don't have to count all the time. You can look at your notes from yesterday to review the methods we discussed. Any ideas? BRIAN?

(b) After editing:

T So by looking at- this one is ... eight square units this one is twelve square units this one's twenty-four square units how do I find the area of a rectangle?

T Now find me a mathematical way of doing it so I don't have to count all the time. You can look at your notes from yesterday to review the methods we discussed. Any ideas? BRIAN?

The above text is organized into two separate turns by determining an appropriate break in the text while keeping each turn within three lines. Appropriate breaks can be made at the end of an utterance that is syntactically and/or intonationally complete.

## 2.6 Time Coding

Once a lesson has been completely transcribed/translated, speaker turns will need to be time coded. It is important that the time coding procedures occur as the *last* step of the second pass, rather than a *simultaneous* step, since it provides a final chance to review the transcript. This is a separate procedure that will be explained in detail during a training. The process involves placing time codes on each turn segment. More details will be provided in the training session.

## 2.7 Logging Hours

There is one other form that should be filled out whenever you begin to transcribe/translate: Transcription/Translation Log. This form is used to keep an accurate record of the amount of time it takes you to transcribe or translate each lesson. Whenever you sit down to work, make a note of your beginning time and the time you finish working. This information will be used to help us predict how many hours it will take to transcribe or translate each lesson. This information is crucial for establishing our project timeline and budget allocations.

Both forms, "Glossary/Problems" and the "Translation/Transcription Log", should be kept together in a single file folder labeled with the name code of the lesson you have been working on (e.g., US-0004). This file should be kept in the filing cabinet in the main transcription area. You will be requested to turn these forms in along with the lesson folder once you have finished transcription/translation.

## 2.8 Completing a Lesson

The second pass of the transcription/translation is finished when you have completely reviewed the entire lesson. If there are sections of the lesson that you cannot transcribe or translate, this matter should be brought up in an individual meeting with the transcription supervisors. All issues noted on the Glossary/Problems form should be resolved before you move on to the next tape. *Please make sure you print out a hard copy of the transcript for the lesson file.*

- Use checklist to make sure all forms are complete
- Print out transcript
- Put all forms and transcript in lesson folder
- Place lesson folder in "In Box" on file cabinet

## Chapter Three: Transcription/Translation Issues

The discourse of both the teachers and students should be reflected in the transcription as accurately as possible. Given that the study includes videodata from seven countries, the process of transcription will vary depending on the language of the classroom. In the American context, standard spelling conventions of American English will be followed. Regional accents should not be reflected in the transcription, but production problems, such as mispronunciation, are to be reflected in the transcription.

In the other linguistic contexts, translators should attempt to capture, as accurately as possible, the meaning of the original language terms in English, without sacrificing



readability. There will always exist a tension between a literal translation and a readable (i.e., “flowing”) text. Translators must, therefore, make on-going decisions, in consultation with other members of the team and the transcription supervisor, concerning such issues.

### **3.1 T/T Meetings**

Regular group meetings will be scheduled to deal with problematic issues regarding translation and transcription. During these meetings, the translators and the transcription supervisor will discuss and agree upon standard translations for mathematical and non-mathematical terminology and phrases. Translators should use the glossary forms that are provided to record any technical terms used in the data, along with the English translations they use when transcribing. The translators’ glossary will be assembled, continually updated, and distributed to the members of the translation teams *so that consistency and uniformity can be achieved in the transcription/translation of the videotapes*. If members of the translation teams cannot readily agree on succinct translations of terms or phrases, members of the larger mathematical education community will be consulted for their input.

Some dictionaries and mathematics and science textbooks are provided in the main transcription area for your use, though we do request that you return them to their proper place after you have finished using them. If you feel that there is any other resource (or resources) that would make translation easier, please make your request known. We are open to any suggestions and look forward to a highly interactive and highly collaborative approach to the translation of the videodata.