Guidance for Analysing Video Recordings of Undergraduate Supervisions

The utility of analysing video/audio records of supervisions

This guide to recording and analysing supervisions has been prepared summarising the methods used to capture and analyse records of supervisions.

Reasons for doing this can be summarised as:

- The usefulness in raising questions across supervisions comparatively for example 'What is a typical Plant Science rather than a typical Engineering supervision like?'
- The usefulness in raising questions within supervisions.- for example 'What is the balance between lecturer and student contributions?'
- The opportunity to observing a range of practices in otherwise difficult to access situations.
- The opportunity to compare intentions, expectations and beliefs about supervisions with the reality of what is observed.
- The opportunity to use the data for oneself diagnostically. For example 'Can some of the lowest D values taken with measures of low volume of talk be correlated with students who could be considered as needing more support in verbalising their understanding?'

Resource issues

The following four stages are discussed.

- Generating an audio and video record.
- Transcription of the audio record.
- Quantitative analysis to generate a metrics profile.
- Qualitative analysis to generate collections of coded video clips.

The equipment resourcing is largely tied in with the **recording** stage¹. While recording the supervisions is time efficient, preparing the data for analysis and subsequent analysis is time demanding and a cost benefit analysis will be needed to be performed to decide on both sample size and how much of the process as described in this report a Faculty wants to/can dedicate to such research. A reasonable sample might be to consider recording 5 lecturers during 2 of their supervisions.

The **transcription** process falls into two stages. Firstly, the transcripts need to be generated in a way that is compatible with the two analytical systems chosen (see below for recommendations of software). The CLAN software is the most particular in its requirements. The transcriber used in pilot work reported

¹ For these methods refer to the document 'Video Procedures for Supervision Filming'.

A Fox

December 2006

CMI Pedagogy Project: Small group supervisions

needing 5 hours per hour supervision to generate a CLAN ready transcript. Secondly, in addition for use in TRANSANA software, time codes need to be allocated to turns to allow the audio record to align with the video record. This is carried out within TRANSANA and can take a further 2-3 hours per video record.

Analytical packages can be downloaded free of charge (see below) and do not have particular requirements beyond an up-to-date PC/Mac. However, TRANSANA is difficult to output from and the pilot project decided to invest in a portable/external hard drive to hold the data and databases on. A suitable one can be obtained for around $\pounds 60^2$.

Quantitative analysis, as described using CLAN software below, is very quick to run to generate the metrics profiles. About one hour would be needed to work through a single supervision. More time is then needed to explore the data in Excel.

The final stage of **qualitative analysis** is more time consuming and a Faculty would need to reflect on the intended uses of the analyses before deciding to embark on this stage. Once a framework has been entered into a package like TRANSANA i.e. a database (one single one is recommended), data attached as transcripts and videos (episodes) and keywords from the coding frame entered, interesting extracts from each supervision can be selected and dragged into collections. These can be allocated keywords and described in a title and longer comment. The collections can later be searched according to keywords. For each transcript perhaps a day should be allocated for the basic analysis. More time is then needed to explore the data comparatively and in relation to the questions posed by the earlier metrics analysis.

If a Faculty wanted to budget for a set of supervisions to be analysed in terms of human resourcing, the analysis recommended above would therefore require a dedicated researcher to work full time for 6 weeks. This excludes the recording and video/audio record preparation time.

The estimate is calculated on the basis of supervision records from 10 supervisions. **Quantitative** analysis would involve approximately 10 days of researcher time - although this is a minimum estimate and does not allow for training. This figure is made up of 50 person hours of transcription (assuming the transcriber was *au fait* with the requirements of CLAN and how to keep the transcript free of hidden formatting codes) followed by 10 person hours of primary analysis using CLAN (assuming the analyst was up to speed with using CLAN) and 2 days to generate a preliminary report. If **qualitative** analysis is also being considered a further minimum of 16 days of researcher time should be considered. This allows for 25 person hours of transcript preparation and 10 days of preliminary analysis followed by 3 days to generate a preliminary report.

² For example (November 2006):

http://www.misco.co.uk/productinformation/~122078~/product.htm?affiliate=2004 60 Gb, £58.74, no power supply needed. <u>http://www.cclonline.com/product-info-pc.asp?id=7433&pc=kelkoo</u> 40 Gb £49.89, no power supply needed. <u>http://www.it247.com/s.nl/it.A/id.48349/.f</u> 40 Gb £53.22 no power supply needed.

• For quantitative analysis

The transcript of a supervision can be loaded into quantitative analytical software packages. We recommend CLAN software, originally used with the CHILDES project³ (MacWhinney, 2000), which is available to download free from (http://childes.psy.cmu.edu/clan).

• For qualitative analysis

Two packages are recommended.

One possibility is to use the powerful workbench Atlas.ti; downloadable from their website <u>http://www.atlasti.com</u>. This will handle the transcript of the videos more successfully than the video imagery. A free trial version is available from the site. A single user version currently costs $\pm 500^4$.

A further possibility is to use TRANSANA software; downloadable free from their website <u>http://www.transana.org</u>. This user-friendly software allows for analysis of the transcript and image simultaneously and the generation of collections of exemplar clips.

Transcript generation

• For use with CLAN

The transcript needs to be generated in a CLAN-ready format which entails applying particular conventions⁵. The files need to be text files and the software is very sensitive to errors in transcription relating to the start of lines and to hidden formatting codes.

Comment: Add sheet from Patrick on how to get a transcript CLAN-ready

• For loading into TRANSANA

The transcript needs to have been saved as an .rtf file in which each speaker's turn begins a new line which will start with a speaker code e.g. LEC (lecturer) before the speech. The VOB format of the DVD needs to have been converted to Mpeg 1 or 2 format.

³ The acronym CLAN stands for Computerized Language Analysis. It is a program that was designed specifically to analyze data transcribed in the format of the Child Language Data Exchange System (CHILDES). Leonid Spektor at Carnegie Mellon University wrote CLAN and continues to develop it. The current version uses a graphic user interface and runs on both Macintosh and Windows machines.

⁴ Price checked on the website 10.10.2006.

⁵ Refer to PC's 'CHAT transcription system: 1-page overview'.

A Fox December 2006 CMI Pedagogy Project: Small group supervisions

Useful metrics that can be generated for a supervision record

It is suggested that quantitative analysis is a good place to start an exploration of supervision records. Using the CLAN software and three basic functions a metrics summary for each transcript can be generated⁶.

The mlt algorithm generates the following for the whole group:

- number of utterances,
- number of turns,
- number of words,
- ratio of words to turn (how long turns are),
- ratio of utterances per turn (number of sentences per turn),
- ratio of words per utterance (length of sentence).

The freq algorithm, which needs to be run for each individual in the group⁷,

- lists the words used in the group,
- states the number of different words (types),
- the number of total words (tokens),
- the ratio of type to tokens (measure of lexical variety) NB. very dependent on length of talk.

The vocd algorithm, again to be run for each individual, calculates from a sample of their text the change in type:token values and generates a summary D value (which is a better summary measure of lexical variety).

Key features of the metrics profiles generated for a supervision

A measure that affected quantitative differences according to Faculty was supervision group size. This impacted on the number of turns made by each group member. Those that demonstrated differences between lecturers according to Faculty were numbers of sentences (utterances), lengths of sentences (words/utterance), length of turns (words/turn) and numbers of total words. Differences between students according to Faculty were mainly related to length of sentences and length of turns.

Unequivocal measures were utterances per turn for students (all approx. 1 for students) and that students took a similar total proportion of turns compared with the lecturer (just over 50%). Students also generated a higher range of words compared with numbers of words (type:token ratio) compared with their lecturers.

Measures that appeared to vary according to individual rather than Faculty were, for lecturers, the D value and number of sentences per turn. For students most

⁶ The CLAN software initial screen prompts you to attach a folder and then, in the space provided, the abbreviation for algorithms are entered. You are given the opportunity to relate this to a particular transcript and then to press 'run'.

 $^{^{7}}$ In the CLAN software the abbreviation for the computation is entered followed by +t and the individual's code e.g. freq +tLEC (to compute the word counts for the lecturer only)

A Fox December 2006 CMI Pedagogy Project: Small group supervisions measures showed a wide range in variability, regardless of Faculty e.g. number of turns, number of words, D value and sentence length.

Entering the metrics data into an Excel spreadsheet allows exploration of the balance of contributions to be made. The proportions, rather than frequencies, of metrics can be generated. This is interesting to do with respect to percentage of the total for the group e.g. number of turns, number of words and also as a proportion of the lecturer's values. Such secondary analysis might raise questions which can be explored using qualitative methods. These could relate to inter-supervision comparisons or intra-supervision queries. Examples from the pilot work included 'What is about the nature of the [a particular] supervision that means the lecturer is taking a low proportion of the turns and sentences? and 'Why does [a particular] student(s) contribute so little?'

Analysing supervision records using TRANSANA software.

The user interface of TRANSANA

There are four windows within TRANSANA.

- The top right is the *video window*.
- The top left is the *sound window* which displays the waveform (a visual representation of the sound record) and timing information.
- The bottom left is the *transcript window* and
- The bottom right the *database window*. This displays the TRANSANA file management information.

Procedural guidance on getting started

Firstly, a *database ID* needs to be generated. The pilot one was called 'CMI supervisions 0506'.

Within this, *series IDs* can be used to cluster supervision records. A series is a collection of videos. Series ID could be a tutors or Faculty name.

Before adding any videos, **enter a keyword group** for later coding. Double click 'keywords' in the database window to add the keyword group e.g. 'CMI framework' and then click on the created group to add individual keywords. These are clustered into subgroups and will need to be entered as e.g. 'MLE: learning materials'

A Fox December 2006 CMI Pedagogy Project: Small group supervisions Suggested keyword group and keywords

Sub- group cluster	Full Title ⁸	Keywords (and phrases)
MLE	Making learning Explicit	Learning materials Exemplars Learning objectives Links
ст	Contingent Teaching	Diagnostic questioning Common misconceptions Conceptual difficulty Scaffolding
AL	Authentic Learning	Real world Modelling Domain-specific
SR	Self Regulation and Independance	Peer assessment Self assessment Student initiated items Transferable skills
СА	Constructive Alignment, Synthesis and Throughlines	Authentic assessment Assessment aligns with teaching Course elements align Synthesis
ТА	Transparency and Accountability	Explicit criteria Aims and objectives
PE	Praise and empathy	Show concern for student difficulties Make sure workload is appropriate by monitoring students' experience Reward understanding Make formative criticism Lecturer enthuses

In TRANSANA videos, in this case of 1 hour long supervisions, are called *episodes*. **To upload a video**, right click on the series and choose '*Add episode*'. Fill in the ID, naming each by date/supervisor and giving details of those participating, and browse to upload. A form prompts you to add a transcript when you add an episode and assign a default keyword group to the episode. **Upload**

⁸ The warrant for these code clusters is outlined in an Appendix to this guidance and is justified more fully elsewhere on the Camtools website. The above table has been adapted from this earlier work and is justified in the Summary Report: Qualitative Analysis of 2005/2006 Videos of Undergraduate Supervisions from 2 Cambridge Faculties.

A Fox December 2006

CMI Pedagogy Project: Small group supervisions

the associated transcript by browsing as before. These should be rtf forms of the transcript. The *transcript ID* could be the same as the episode ID to avoid confusion. TRANSANA also prompts you to **create a waveform** when a video is loaded.

After loading the transcript, **the transcript needs to be aligned with the video** so that the transcript moves through turn by turn automatically as the video plays. Right -clicking on the waveform starts and stops the video. If you want to add time codes it's easiest to use this control. You must be in *edit mode* to adapt the transcript. To do so, select the button with the '*hand*' in the transcript toolbar. Listen for the beginning of the next speakers turn and right click on the waveform to stop. Move the cursor to the corresponding place in the transcript and press '*Current*' in the sound window. A red mark will appear in the transcript window. Right click on the waveform again to continue. Whilst in edit mode, any other changes to the transcript can be made. **To mark and time pauses** in the commentary, drag the mouse over the flat area of the waveform and press '*Selected*' in the sound window toolbar. It will tell you how long the pause is. If you click on the area in the transcript it will insert the length of the pause in brackets.

Playing a transcribed video

• Using the video window

Straightforward play, pause and stop buttons are available. You can also change position by dragging the video tracker and fine tuning with the mouse wheel. The transcript should scroll to follow the video if time codes have been inserted.

• Using transcript window

Left click on a spot in the transcript, then right click. The video will play from time code before the cursor to the one after. To play a larger passage highlight it and right click.

• Using the sound window

Click on waveform positions and then right click to play or pause. Drag to select a segment. You can zoom into waveform for greater precision.

• Using keyboard shortcuts in the transcript window

- Ctrl S To start or stop video, it will automatically rewind 2 seconds
- Ctrl D start and stop with no rewind

Ctrl A - Rewind 10 seconds

Ctrl F - Fast Forward

A Fox

December 2006

CMI Pedagogy Project: Small group supervisions

Procedural advice on analysis

Clips (extracts of video image and associated transcript) can be generated and keywords attached to those clips. These clips can be sorted into groups of analytically important *collections*. Double click 'Collections' in database window **to create a collection**.

Collection (Cluster)Full Title
MLE	Making learning Explicit
СТ	Contingent Teaching
AL	Authentic Learning
SR	Self Regulation and Independance
CA	Constructive Alignment, Synthesis and Throughlines
TA	Transparency and Accountability
PE	Praise and empathy

Suggested collection names

To make a clip, the transcript must be in edit mode. Highlight the section of the transcript of interest, being careful not to include the time codes at the beginning or end of the section. Click on the highlighted text and drag to the appropriate *'Collection'* in the database window. Give the clip an ID relating to the tutor and the key evidence. Automatically it will be given the keyword group assigned to that episode (video). You can choose to add more keywords from an existing list or add new ones. **Select the keyword(s) most appropriate to that clip** by clicking on those on the left list and see them appear on the right.

To search for keywords, right click on 'Search results' in the database window. Give the search a meaningful name. You can use a Boolean search e.g. and, or, not etc. Once the search is completed the results can be converted into a new collection, saving the results as a collection of related clips. Searchs cannot be carried on lecturer.

To generate reports, left click on any collection you have created to get printer ready reports. Useful reports are the *Collection Summary Report* (all the clip IDs, start and stops times, duration and transcripts) and the *Keyword Usage Report* (a list of keywords for all groups in a collection).

Appendix: The warrant for and emerging interpretation of analytical themes

The items and operationalisations below are not exclusive lists and there are likely to be other activities which fit in under these broad categories.

Theme	Key word cluster	Example Items/Operationalisations	Warrant/background	
Making Learning Explicit	MLE	 Availability of appropriate learning materials Exemplars of different standards Learning objectives shared and referred back to Students encouraged to identify links between course elements 	ETL, AFL, TfU	Comment: Need to expand on these and give exemplar references
Contingent Teaching	СТ	 Diagnostic questioning Addressing common misconceptions Polling for areas of conceptual difficulty and responding Scaffolding explanations and problems 	Constructivism, Vygotsky, Ausubel, Wood	
Authentic Learning	AL	 Students encouraged to relate concepts to 'real world' examples Modelling ways of thinking and practicing Encouraging students to use domain-specific 'ways of thinking and practicing' 	ETL, Biggs, WTP, Constructivism	
Self Regulation and Independence	SR	 Students supported in peer assessment Students supported in self assessment and personal goal setting Student generated activities Development of transferable skills 	Constructivism, AFL, Self- efficacy research	
Constructive Alignment, Synthesis and Throughlines	CĀ	 Assessment aligns with authentic ways of thinking and practicing Assessment aligns with teaching Course elements align Synthesis of content from 	QAA, ETL, Biggs	

A Fox December 2006 CMI Pedagogy Project: Small group supervisions

		different course elements	
Transparency and Accountability	ТА	 Assessment criteria and processes made explicit to students and supervisors Course aims and objectives made explicit to students and supervisors 	Fielding?
Praise and Empathy	PE	 Show concern Ensure workload is appropriate by monitoring students' experience Reward understanding Make personalised formative criticism Enthuse 	AFL, personalised learning, ?

It is recognized that some of these themes might be in tension with each other for example, an overly prescriptive approach (high on MLE) might equally lead to low levels of student self-regulation (low on SR).