DOCUMENT RESUME

ED 148 325 IR 005 368

AUTHOR McAleese, Ray
TITLE The Systematic Observation of Lecturing--Its Use in Training and Research.
PUB DATE Sep 73
NOTE 19p.; Paper presented at a seminar on Classroom Research at the University of Leicester (Leicester, England, September 26-28, 1973); For related documents, see IR 005 366 and 369

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Classroom Observation Techniques; *Closed Circuit Television; Content Analysis; *Lesson Observation Criteria; *Microteaching; Self Evaluation; Teacher Attitudes; *Teacher Behavior; Teacher Education; Teacher Evaluation; Universities

ABSTRACT

A seminar is described in which systematic observation techniques were applied to lecturer training to make the supervisory element in microteaching more objective. The system provided a highly structured technique that the lecturers could use to systematise their discussion of microlectures or real lectures, either recorded on tape or live. Three parameters were "observed": non-verbal, verbal, and content aspects of lecturing. These categories were used after the trainees learned the system and could reliably code a lecture into these components. This information was then used by the trainer and trainee as data in analyzing a lecture. A small scale study is reported in which the attitude of the trainees to each of four treatments (systematic observation vs. conventional feedback and audio vs. closed circuit television augmented feedback) was measured. No significant difference was found between groups. A discussion of the reliability and validity of the system is appended. (DAG)

The Systematic Observation of Lecturing - its use in training and research.

Ray McAleese

Department of Education,
King's College,
University of Aberdeen.

September, 1973

A paper presented at the SSRC seminar on Classroom Research held in the University of Leicester, 26-28 September, 1973.

The Systematic Observation of Teaching

My contribution to this seminar will be concerned with the work that I have undertaken in the University of Aberdeen. Conventional microteaching techniques have been successfully employed. In this work, however, I have attempted to see if I could apply systematic observation techniques to lecturer training.

In Dublin, I used systematic observation (FIAC categories) to make more objective the supervisory element in microteaching training. (McAleese, 1973) I have found that the FIAC-type categories provided an objective framework for supervisory sessions. I was therefore very keen to see if the same approach could be taken with university lecturers. To begin with I decided to concentrate on lecturing (didactic or monologue teaching) - after 15 months I have some hopes of increasing the coverage of the technique to tutorials and practicals, but in the first instance the work has been concerned with lecturing. This decision to concentrate on lectures was taken primarily on grounds of convenience. The attraction of systematic observation to tutorials was outweighed by the fact that even in 1973 lecturers still find most of their teaching time in formal lectures! In practice, lecturing is still, to those involved in training, the more important activity (a point which may not be innovative, but reflects the aspect of the trainers role concerned with maintaining the status quo.) Although initially I was hopeful that Simon and Boyer (1970) could give me some leads to

possible observation systems, this was not possible. I had to start from scratch in the construction, validation and implementation of a new category system. Lecturing and lecturers I discovered were different in several ways from either what is irreverently called 'ping-pong' teaching or from any other type of teaching analysed in Mirrors.

Firstly, lecturing is basically non-interactive: it is monologue. Although this may or may not be desirable depending on one's viewpoint, lecturing is not 66% teacher talk (e.g. secondary 'ping-pong') but 100% teacher talk! Therefore, any system of interaction analysis had to be discounted.

Secondly, the teachers that I was to be dealing with were by and large blissfully ignorant of educational jargon! You might say that this is an advantage, but it does pose problems when one thinks of using same system that is based on instructional concepts inbedded in educational phraseology, e.g. Taba or Aschner and Gallagher. (Taba, Levine and Freedman, 1967; Aschner and Gallagher, 1965).

Thirdly, I required a system that would be applicable to all the various disciplines in the university - from medicine to mathematics. A requirement that may need modification in light of the diversity in style shown by some of my colleagues in various subjects that we have looked at.

It was decided that the system should provide a technique that the lecturers could use to systematise their discussions of microlectures or real lectures either recorded on tape or live. Using the Simon and Boyer (1970) classification, I was looking for the system that could be described by the following

parameters. It can be:-

1. Used on any subject matter.
2. The sole subject of observation is the lecturer.
3. Only one lecturer need be observed at any one time.
4. It should be able to be used in research, training and evaluation.
5. Data must be able to be collected live, using one observer.
6. The coding unit should be time units.

I decided using these parameters to 'observe' the pedagogic structure of the lecture. I had the choice of looking at what might be called either the Non-verbal, the Verbal or the Content aspects of lecturing. Verbal structure was chosen in the first instance as it provided fewest problems given the constraints of didactic teaching and the nature of the lecturers.

The analysis of verbal structure was also chosen as it provided what can be called a meta-language of supervision. As most of the lecturer training in Aberdeen is involved with heterogeneous groups of lecturers, a problem is that one has in the same course mathematicians, chemists, geologists, etc. The variety of discipline represented is by and large an advantage, but when it comes to discussing the individual content of microlectures the trainer is presented with something of a problem. The participants, teaching their own subject, find that comparisons between lectures is also very difficult. As this is one of the key elements of the training - that is making peer comparisons, a system that

alleviates this problem is to be preferred. By analysing the content of lectures in terms of Teaching Points, Examples, Asides, etc. it is possible to construct a very simple meta-language of instruction.

Figure 1 about here

The categories in Figure 1 are the words of this language and also the system of observation itself. The participants are able to communicate their value judgments about their lectures by using this language, instead of using mathematical, chemical or geological terms. Taking this metaphor just a little further I hope that it may be possible to construct a 'grammar of lecturing,' although this is not possible at present. The construction of the grammar can be thought of as the research element of this ongoing study.

Having now described how I came to use the system and having given you some idea of its application let me say how I have used the system in training and research, (a note on the reliability and validity of the system is contained in the Appendix to this paper).

TRAINING

One reason for using a systematic observation system is to give a high degree of structure to the supervisory session in the microteaching cycle. Normally when the trainees have viewed their microlecture the supervisor adopts a non-directive role and using peer criticism to emphasise points - he "shapes" in a non-threatening way the behaviour of the trainee.

Sessions like this, no matter how good the supervisor is, can become rather rambling and more non-directive than desired! Systematic observation data gives a focus to the discussion and also provides data that describe the teaching performance in terms that have unique meanings. Before the category system can be used the trainees must learn the system and this requires an initial training session that is devoted to shaping the trainees' behaviour in recognising and recording the various categories. Figure 2 summarises this process.

Figure 2 about here

Having completed the training, the lecturers should be able to listen to a lecture and reliably code the content on to a Time-Line Display (TLD). The TLD is used as it permits quick inspection of the lecture and keeps the data in its original sequence. Further, the TLD allows the trainees to look for what we are calling the 'profile of the lecture'. That is the number of "ups" and"downs", etc. that can be seen in category changes.

Figure 3 about here

During the actual training sessions this information is used by trainer and trainee as data. That is, if the trainee wants to make a point about a certain aspect of the lecture, he can refer to the TLD for an 'objective' evaluation or record of the lecture at that point. He can use the data to facilitate explanations of why any part of the lecture was less good than other parts. The trainee can refer back to a particular point

in the lecture and find out whether the content analysis can help explain why the lecture seemed to go wrong. This type of analysis begs for 'Norms' to be established for different types of lecturing. It is this stage of the research that we are now moving on to as we intend to undertake a study of lecturing in Aberdeen and apply these categories to a variety of lectures in an attempt to establish some idea of 'Norms'. This work has just started and so far we have no usable results. We are undertaking a correlational study and we hope to look for relationships between lecturer verbal behaviour, (teaching effectiveness?) and student performance. It is also our intention to take the correlational study further by establishing what may be causal relationships in experimental studies.

Let me conclude this contribution by reporting on a small scale study that I undertook earlier this year in an attempt to investigate how this systematic observation was received by lecturers in training. I used a 2 x 2 design for the independent variables. Variable A ..... Systematic observation vs. Conventional feedback. Variable B ..... Audio only augmented feedback vs. CCTV augmented feedback. The dependent variable was the attitude of the trainees to each of the four treatments. Not a very 'strong' variable but one that we feel is important if this developmental work is to succeed. Using multiple choice and open-ended questions for the dependent variable there was no significant difference between groups. A result that was in one sense not unexpected due to the crudity of the dependent variable. However, other explanations of the non-significant differences that are of more immediate

interest to my work would be that there was a large Hawthorne effect on our trainees. However the open-ended questions indicated that the trainees did accept this new type of training and we are further encouraged by follow-up that the lecturers can do themselves. I think that this follow-up is particularly attractive as after the initial training is over, as the lecturer applies the categories to their own lecturing in vivo and attempts to find behavioural explanations of successful or 'disastrous' lectures. This work is still in embryo.

My research and development has just started: it will I hope be possible in another year or so to make a more objective evaluation of its efficacy. At present we have still a lot of problems to solve. For example, how to measure change in behaviour, how to make statistical comparisons between 'profiles'. Looking back to the early sixties I think that this work is in about the same position as the observation systems were in about 1962 or 1963. I am, however, hopeful of similarly successful termination of its gestation period!

APPENDIX

The Reliability and Validity of the Categories

1. The starting point of the system is the research reported by Gage (1968). In this he suggested that one element of lecturing that might discriminate between effective and ineffective lecturers was their ability to relate examples (Egs) to teaching points (Rules). This 'rul-eg' unit, he suggested, is found more in effective than in ineffective lecturers. Rules (Points) and Examples give the first two categories of the system.

2. The other categories were devised by systematically observing lecturing and deciding on recognisable events that make up the rest of this didactic teaching (McAleese, 1972). In initial trials it was found that lecturing could be very crudely divided into TEACHING (Points (P), Examples (E)) and MANAGEMENT. The management section could be sub-divided into two categories, Asides, that were instructional in intent (Ai) and Asides that were non-instructional, or useless (Au),..... in relation to the teaching of the topic. The fifth category X was added as it was decided to analyse the content using a time unit method of sampling and a dustbin category was required.

3. Analysis of the sentence structure of lectures indicated that on average the time taken to say one recognisable 'thought unit' (Taba, 1967) .... e.g. Point, Example was 11.00 seconds. Using this piece of information, it was

decided to adopt a 6 second time unit for sampling, in training at least. Some other work using 3 second intervals has been attempted as the shorter time unit causes fewer decisions about multiple coding during any particular time unit. The shorter sampling is also more applicable to research and validation. The validity of this point is still to be established. If the mean time of a thought unit is greater than the length of coding time unit, then there may be problems in using conventional I-A statistics.

4. The five categories were applied on two occasions to samples of teaching with the following results. Clearly, Examples are the most difficult to operationalise and therefore to observe.

Table 1 about here

This gave a measure of the Construct reliability or Stability of the system.

5. A measure of the concurrent validity of the system was attempted by devising a 12 point rating scale instrument that paralleled the observation system and applying this to 18 samples of lecturing. A test-retest reliability coefficient for the rating scale was estimated as 0.8 Correlations for each of the categories with the rating instrument are found in Table 2.

Table 2 about here

6. Further and more detailed validation of the system is underway.

DEPARTMENT OF EDUCATION
UNIVERSITY OF ABERDEEN

CONTENT ANALYSIS OF LECTURING (C.A.L.)

Glossary	Coding Unit	Symbol
A teaching point, part of an argument or thesis. An item of information, an opinion, a fact, a generalization, a principle or concept. Lecturer's own opinion or quoting another authority.	POINT	P
Examples, related to teaching points. Comparitors, a concrete example of a teaching point.	EXAMPLE	E
Verbal behaviour of relevance to the lecture but not intended as part of an argument or thesis. References, objectives or aims - relevant to the lecture or course of lectures. Summarizing teaching points made in other lectures.	ASIDE (Instructional)	Ai
A non-instructional aside. Comments of no direct relevance to topic of lecture. Greetings, tension releasers, directions, jokes or irrelevent examples.	ASIDE (Non-instructional /useless)	Au
Verbal behaviour that is uncodable. Confusion or silence. Non-verbal communication. Writing on blackboard, using A-V aid; Student Talk.		X

Figure 1 : The Categories.

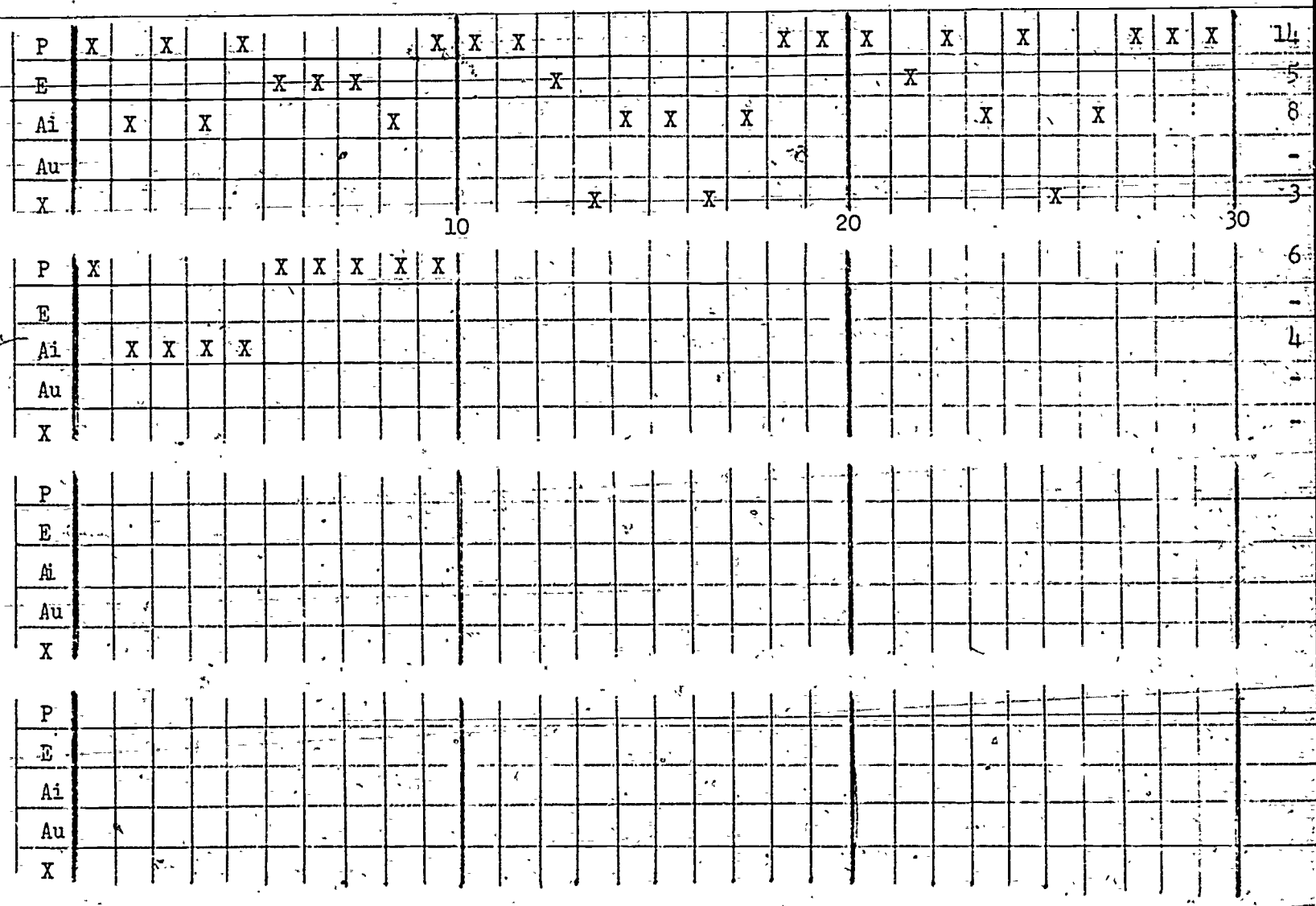
DAY 1	DAY 2
Training in Observation System	Training Session
1. View video tape of lecturer in lecture theatre.	1. Videotape Lecturers 1 to 4 - 9 minute micro-lectures
2. Discuss	2. Code first 5 minutes lectures, 1 and 2. (as a group)
3. The categories of the observation system	3. Code last 5 minutes lectures, 3 and 4 (as a group)
4. Listen to examples of categories	4. Replay lecture 1
5. Listen to lecture (same as 1), looking for examples of categories	5. Discuss, with codings
6. Discuss	6. Replay lecture 2
7. Code transcription of lectures (same as 1)	7. Discuss, with codings
8. The Time-line display (TLD)	8. Replay lecture 3
9. Transfer coding of lecture (7.) to TLD	9. Discuss, with codings
10. Sampling of teaching behaviour (3 seconds; 6 seconds)	10. Replay lecture 4
11. Code Typescript of lecture (same as 1) at 6 second intervals	11. Discuss, with codings
12. Transfer typescript (11.) to TLD	12. Evaluate the recordings and the codings
13. Sampling at 6 second intervals, (listen to timing)	
14. Listen to lecture (same as 1)	
15. Code lecture at 6 second intervals (same as 1.)	
16. Transfer lecture (15.) to TLD	
17. Discuss	

Figure 2: Training Programme

UNIVERSITY OF ABERDEEN
DEPARTMENT OF EDUCATION

UTM/73/0033

Coded at 6 sec. intervals



SUMMARY		%
P	20	50%
E	5	12.5%
Ai	12	33.4%
Au	-	-
X	3	4.1%
	Σ 40	

Coder

R.McA

Lecturer

Mr. James

Date

25/4/73

th

mins. .... secs.

NOTES:

16

Figure 3: Time Line Display Profile for Lecture

Category	Percentage Agreement	N = 29
P	95%	
E	57%	
Ai	98%	
Au	93%	
X	97%	

Table 1:
Stability Coefficients for the Categories
(ten days between two observations, one trained observer)

Category	Product Moment Correlation
P	.51
E	.82
P + E	.54
Ai	.75
Au	n.a.
Ai + Au	.75

N = 18

12 point rating scale

(n.a. not enough data)

Table 2:

Concurrent Validity of the Categories

REFERENCES

Aschner, M.J., Gallagher, J.J. et alia (1965) Aschner - Gallagher System for Classifying Thought Processes in Context of Classroom Verbal Interaction, Institute for Research on Exceptional Children, University of Illinois, U.S.A.

Gage, N.L. (1968) An Analytical Approach to Research on Instructional Methods, Journal of Experimental Education, 37, 119-125.

McAleese, R. (1972) An Analysis of Microteaching Videotapes Made During a Teaching Methods Course for University Teachers, University of Aberdeen (UTM/72/005).

McAleese, R. (1973) Interaction Analysis and Teacher Training in Trinity College, Ireland, Classroom Interaction Newsletter, 8, 2,

Simon, A. and Boyer, E.G. (1967) Mirrors for Behaviour, An Anthology of Classroom Observation Instruments, Research for Better Schools, Inc., Philadelphia, U.S.A.

Taba, H; Levine, S. and Freeman, E. (1967) A Scaled Analysis of Performance in Classroom Discussion of Three Cognitive Levels, Paper presented at the American Educational Research Association Convention, Chicago, U.S.A.