

Exploring CertTESOL tutors' beliefs on effective teaching

Catherine Rosenberg

Abstract

The aim of the study outlined in this article was to examine six teacher trainers' assessments of effective teaching in relation to each other and the reasons and justifications they gave for their judgements. A questionnaire was devised in which they were asked to rate and rank a range of criteria, including any they wished to add. This was followed by interviews in which the findings from individual tutors' questionnaires were discussed with them. Inspection of the data revealed that when there was agreement about the importance of criteria, then the reasons for these beliefs had much in common; however, such agreement was not predominant.

Background to the study

Observation of trainees' teaching is an integral and unavoidable part of the assessment of teachers on the Trinity CertTESOL¹ run by this university faculty. A key issue is the level of consistency in the feedback and advice we give to our trainee teachers: the existence of inconsistency between observers may be perceived as undermining the fairness and validity of the assessment process. While attempts have been made to qualify what exactly constitutes effective teaching, many researchers reject the possibility of arriving at any universal consensus. This in turn renders the evaluation of teaching, a necessity on teacher training courses and in terms of quality control in the teaching profession, more problematic than it might superficially appear.

The three studies that were seminal to the research described in this article were those carried out by Cook and Richards (1972), Mackay (1989) and Piper (1996). Cook and Richards (1972) carried out research whereby teachers were independently rated by their principals and supervisors on scales of teacher competence. When the rating scales were intercorrelated, there was a high level of clustering which showed the degree to which principals and

¹ Trinity College London Certificate in Teaching English to Speakers of Other Languages

supervisors were in general agreement about their ratings. The researchers conclude that information about a teacher's performance based on rating scales should be interpreted with caution, as such scales say more about the rater than about the teaching they observe.

Mackay (1989) dismisses the viability of any global concept of effective teaching, and suggests that assessment of teaching should include an element of negotiation between all interested parties. His research, carried out in relation to the RSA Diploma of that time, concludes that different assessors do not consider the same features of classroom performance to have the same degree of importance in their contribution to successful teaching, and that assessment may be as much a feature of the assessor as of the candidate and the way in which the lesson proceeds.

Piper (1996) built on this research in relation to the RSA/UCLES CTEFLA. Her conclusion is that despite UCLES providing some guidelines for the approval and conduct of CTEFLA courses in order to ensure a minimum of standardisation, tutors do not share a common set of priorities and there are differences between their espoused views and their actual practices.

Following these findings, the aim of this descriptive study was to explore individual tutors' assessments of effective teaching and the reasons and justifications they gave for their judgements. The rationale for the study had two dimensions. The first was to test the findings from Cook and Richards (1972), Mackay (1989) and Piper (1996) in relation to the team of tutors teaching on the Trinity CertTESOL courses at this university. The second was to contribute to the professional development of the team, which it was intended would occur even if few discrepancies were found, because awareness of our own assumptions in assessment would be raised; in the same vein, there could conceivably be a benefit for future trainees who might receive more consistent and critically aware assessment.

In his paper considering how research into teachers' theories in English language teaching can enhance the profession's understanding of instruction, Borg (1999) draws attention to the impact of teachers' theories on their work, and to how teachers becoming aware of their personal theories is central to their growth as professionals. This is in keeping with the developing preoccupation in the literature with teacher thinking and reflective practice. Apelgren (2000), in her study of EFL teachers' theories and experiences of change, makes the point that

More research is needed in the area of teacher thinking to further understand teachers' and learners' constructs and theories underpinning actions, and consequently to understand better how to work for implementation of reforms and professional development.

(Apelgren 2000:2).

A further dimension to this study was therefore its scope to fit in with the need for further qualitative research on teacher thinking in relation to professional development.

Research Design

The six participants in this study were all current tutors with significant contributions to the university's Trinity CertTESOL course. Between them, they have been responsible for virtually all the input sessions and teaching practice observations. Every person took part in the study voluntarily. While the study was small scale, the fact that all six tutors participated gave sufficient numbers for a meaningful comparative study to be made.

A questionnaire (see Appendix 1) was designed both to test the findings of Mackay (1989) and Piper (1996) that there is inconsistency between assessors of teaching practice on teacher training courses, and to serve as a springboard for the subsequent interviews. An objective checklist of criteria for assessing teaching, taken from Trinity's assessment guidelines, was incorporated into the questionnaire, in part following those of Mackay (1989) and Piper (1996). Participants were asked to rate and rank the criteria, adding any further as they wished.

My interest in the second part of the data collection process lay in putting "flesh on the bones" of the completed questionnaires; in using the interview process to find out more about the thinking of each member of the team in relation to the question which is central to us all - what constitutes effective teaching of EFL? My aim at this point was to use the findings from the questionnaires as vehicles for discussion, and to interview tutors individually about their reasoning in relation to their own completed questionnaire. As a data collection instrument, the interview lends itself to asking for reasons *why*, and the qualitative data produced is rendered all the richer by the 'manner in which the response is made' (Weir and Roberts 1994:143).

In this study, the interviews were intended to be complementary to the questionnaires, but were not designed to be used explicitly for triangulation purposes. While a logical consistency between the two sources was to be expected, I was not looking for a perfect match between the quantitative and narrative data. There is an interactive relationship between the two sets of data, in that the results of the questionnaire inform the interview, and I anticipated each participant's interview data would "tell the same story" as their questionnaire data, but the very different natures of the two types of data generated made it unfeasible to make closer comparisons.

The type of interview I chose for this study was what is alternatively known as the 'focused' (Weir and Roberts 1994:145) or the 'semi-structured' (Denscombe 1998:113) interview. It involves the interviewer having an agenda for the interview which is followed on each occasion, but the questions are often open and space is given for individual developments. The advantage to using this interview format is that everyone is asked the same questions (see Appendix 2), making it easier to analyse and compare the data.

In theory, the semi-structured format should ensure that interviewer bias is minimised, although the open elicitation means that it will not be eliminated, as 'bias can ... arise from the interviewer's responses to answers' (Weir and Roberts 1994:143). An additional interviewer effect that is difficult to identify or address is that of interviewees saying what they think I want to hear as oppose to what they really think. All I could do was to try to minimise the likelihood of this through assurances at the beginning of the interview that the procedure was a non-judgemental one, that I was looking for honest opinions, and that there was no "right answer" to the questions I would be asking. In keeping with qualitative research methods as described by Bogdan and Taylor (1975), every participant in the study was viewed as equal, regardless of their standing in the faculty's hierarchy.

Data analysis

Also writing on qualitative researching, Mason (1996:109) describes an interpretive reading of data as involving 'constructing or documenting a version of what you think the data mean or represent, or what you think you can infer from them' and there was a real sense in which the data in this study were read interpretively at the iterative stages of reduction, classification and display. Interpretivism itself, in which human activity is seen as 'text', or as a collection

of symbols expressing layers of meaning (Dilthey 1977, cited in Miles and Huberman 1994) is concerned with ‘how the social world is interpreted, understood, experienced or produced’ (Mason 1996:4). These ideas are central to this work, in that I adopted a qualitative interpretivist methodology in which I attempted to represent participants’ thinking by using an inductive, or bottom-up, approach to analysis.

Questionnaires

Unlike the studies of Mackay (1989) and Piper (1996), this research did not have as its focus large numbers of participants, and it was therefore not relevant to subject the findings of the questionnaire to detailed quantitative analysis procedures. Six questionnaires produced a small data set, and it was sufficient to present the data visually and then interpret them on the basis of simple inspection.

The questionnaire asked respondents both to rate and rank criteria relevant to effective teaching at Certificate level. I decided, however, not to make use of all of the data produced, but to focus in detail on what participants felt to be most important, as this seemed to render the richest data in the interviews. Added to this, I felt that to analyse the data from too many angles would overcomplicate the results, and that conclusions about consistency and individuals’ reasoning could be drawn without having to examine everything that was expressed.

The first step in the analysis of the questionnaires was to provide an overview of the ratings in Sections A and B in order to examine the level of consistency between participants and to see at a glance how individuals were using the rating system. This included all the criteria which participants felt important enough to rate as 5, including those they added themselves.

The next stage was to narrow the focus by viewing those criteria which individuals prioritised as their three most important features of classroom performance in Section C of the questionnaires. This procedure proved the most stark in displaying the level of commonality between tutors. In addition, their three priorities might be found to indicate different “superordinate” beliefs which are likely to affect their judgements.

Interviews

Maintaining this slant, the subsequent stage was to examine the qualitative data produced in the interviews with a view to extracting the reasoning behind participants' top three choices. This process forced me to look closely at what was said and to dissect it meaningfully. In other words, what I was doing was describing and explaining the data: firstly, 'making complicated things understandable by reducing them to their component parts', followed by 'showing how their component parts fit together according to some rules' (Bernard 1988, cited in Miles and Huberman 1994:90).

The first step was to listen again to the tapescripts to refamiliarise myself with the tone and feel of the interviews. The transcripts were then indexed cross-sectionally (Mason 1996); that is to say that each of the six transcripts was examined to find what individuals said on the subject of the three criteria which they had prioritised in the questionnaire. Sections I deemed relevant were then extracted verbatim and placed in a large table systematically. Gordon and Langmaid (1988, cited in Catterall and Maclaran 1997) identify this procedure as the 'large-sheet-of-paper approach', more appropriate for a small data set such as this. I decided to keep extracts in their original form at this stage so as not to confuse the procedure and risk losing the authenticity of the extracts before I was ready to look at them together.

Following the classification, the final stage in the process of displaying the interview data was reduction. The aim of this stage was to narrow the data down to show the reasoning behind the choices made, and in this way reveal something about the personal constructs of each participant. This was the most interpretive of the analysis processes to date, and involved a fair degree of judgement as to what was relevant as well as having to summarise what I thought participants were saying. Sufficient reduction of the data meant that they could be displayed in a matrix following those of Miles and Huberman (1994), which also echoed the table displaying the questionnaire data about the three most important features of classroom performance.

Results

Display of questionnaire data

Appendix 3 gives an overview of the consistency between participants in terms of those features of classroom performance they rated as 5 (the highest possible rating) in the questionnaire. Included in this table are those criteria given in Section A of the questionnaire

which one or more participant rated as 5, as well as any which were given by individuals in Section B and rated as 5. These additions appear in the table following the criterion of ‘Using role play’. It should also be pointed out that those criteria which appeared in Section C as participants’ three most important criteria are included in the table, as not surprisingly they were all rated as 5. The descriptions of the criteria have been summarised from what appeared on the questionnaires in order to economise on space.

Appendix 4 narrows the focus by displaying the consistency between participants in terms of what they ranked as the three most important features of classroom performance.

Display of interview data

The raw data from the interview transcriptions were subject to cross-sectional indexing, which had as its basis the reasons given by participants for their selection of the three most important features of classroom performance in their questionnaires. Further refining of these selected data then took place until it was possible to produce a summarising display matrix, seen in Appendix 5. This matrix incorporates both aspects of the study, in that it illustrates visually the consistency between participants with regard to their three most important criteria, as well as displaying the reasons for their judgements. It also allows for each participant’s constructs to be viewed separately as well as case by case comparisons to be made.

Interpretation of results

Use of rating system

One of the most striking features of Appendix 3, ‘All features of classroom performance which participants rated as 5’, is the different ways in which participants used the rating system. The number of occasions on which individuals gave a rating of 5 ranged from six in the cases of Participants 2 and 4 to twenty in that of Participant 1. This is a useful finding in itself, as it begs the question of whether rating systems can produce distortions of their own. This caveat aside, the table is useful for providing an overview of the consistency between participants.

Unanimity

The most notable observation is that there is only one feature of classroom performance on which there was unanimous agreement as to its high level of importance - that of ‘Aims’, or, in the original questionnaire, ‘Preparation of appropriate aims for a lesson and means of

achieving them'. None of the other 37 categories was deemed sufficiently important by all of the tutors to merit being rated as 5 by everyone, although the following four categories were rated as five by all but one of the participants: 'Establishing and maintaining rapport', 'Classroom management', 'Balancing teacher-student participation' and 'Instructions'. In all but one of these cases it is Participant 2 who does not give the categories a rating of 5, which could raise the question of whether this participant is more conservative in their use of the rating system. Such a question goes beyond the scope of this study, however, and for the purposes of interpretation of what is presented here, we must take the findings at face value.

Differences

An interesting angle to take in interpreting the data presented in this table is to look at the lack of consistency between participants in rating categories as 'very important', as it can be seen at a glance that there are many more points on which we disagree than on which we unanimously, or almost unanimously, agree. There are actually twenty-five categories which are rated as 5 by only one person, with different participants giving the high rating on different occasions. However, this result is not as extreme as it may appear on first reading when one takes into account that on nineteen of these occasions participants are rating criteria which they added in Section B of the questionnaire, and which therefore may not even have been thought of by others in the study. Indeed, this is very likely in the case of Participant 1, who, in filling out this section, did not in every case give 'skills and attributes', as were asked for, but included some activities. This said, there remain six categories from Section A of the questionnaire rated as 5 by only one person, which still far outweighs the occasions when there is complete agreement between participants.

Turning now to Appendix 4, 'Three most important features of classroom performance', the findings here echo those of the previous table in that they too offer a stark representation of the lack of consistency between tutors in their judgements of what makes effective teaching at Certificate level. The table narrows the focus to those features of classroom performance which tutors ranked as their three most important, and interestingly again it is 'Aims' which is the only category to be included unanimously. Clearly this is an area of great importance to everyone, but it remains to be seen whether this is for the same reasons. There was only one other criterion ranked this highly by more than one person, and that was 'Confidence building', selected by three participants. Other than that, a considerable level of inconsistency may be noted between participants, with the remaining nine criteria being selected by only

one person. This reflects a wide disparity in beliefs, particularly as only four of these nine criteria are among those given by participants themselves in Section B of the questionnaire. The argument given above that participants may not have even thought of criteria given by individuals cannot be used with such force here.

It makes sense to examine the data displayed in Appendix 5 horizontally in the case of 'Aims' and 'Confidence building', as we can then see what there is in common in participants' reasoning, and vertically in the cases of individual participants, to see how their judgements form a system and to illustrate contrasting differences in beliefs. Because the same steps have been followed to elicit the data it is defensible to make case by case comparisons.

Appendix 6 illustrates how the reasoning of individuals overlaps and the thoughts they have in common on the subject of 'Aims'. The points in the third column are taken from Appendix 5 and rearranged according to common ideas, the interpretations of which appear in the first column.

This interpretation of the data shows that there is a considerable level of overlap in tutors' thinking on this, the only criterion to have been ranked by everyone as one of their top three choices. There remain five points which were not in common with any others, and so do not appear in the table, but even with this in mind, it can be said that there is more consensus than disparity of opinion on this subject.

An even higher level of commonality can be seen in the reasoning of the three participants who selected 'Confidence building' as one of their top three criteria, as can be seen in Appendix 7.

In this case, the justifications offered by participants could all be categorised within the table, showing that each person's reflections on the subject had something in common with those of at least one other person. What can be surmised from these two tables is that when tutors share beliefs in the ranking of criteria then there is a high level of commonality in the reasons why they believe these criteria to be so important.

Individual systems of thinking

Let us turn now to individual participants and what can be noted regarding their systems of thinking. These can be seen in Appendix 5 by following each participant's column vertically and referring back to the first column for the criteria headings.

Systems of thinking: a discussion

In short, it can be seen that in outlining the reasoning behind their ranking of the three most important criteria, each tutor holds a different set of beliefs. There may be points in common, but the overall impression is of six individual systems of thought. This finding is in keeping with the idea expressed in personal construct psychology that constructs are hierarchically organised, and that the links or associations a person makes between constructs, and how far they are superordinate ideas, reflect the personal meaning they give that construct. A construct system can be described in terms of content and structure, whereby 'the content is the nature of the constructs, the structure the way they are related to each other as a system' (Roberts 1998:31). The findings of a study conducted by Sendan and Roberts (1998) suggest that there is a degree of stability in student teachers' personal theories, but significant changes occur in their structure. Such a distinction means that development can take place even if the content remains unchanged, as individuals change the links between ideas (Roberts 1998). The concept of personal systems of thought developed in this study is consistent with this notion of structure in personal construct systems.

Relationship with findings of other studies

Mackay (1989) and Piper (1996) used much larger sample sizes in their studies into the reliability of assessment of teaching performance at Diploma and Certificate levels. In both cases, their statistical analysis using the Spearman rank correlation coefficient and frequency distribution graphs suggested divergence of opinion, and that tutors do not share a common set of priorities in assessing and developing teaching. As Mackay writes,

... assessment may well be as much a feature of the assessor as of the candidate and the way in which the lesson proceeds.

(Mackay 1989:iii)

While this study is not a quantitative one in which large sample sizes are analysed statistically, the findings resulting from simple inspection of the presented questionnaire data

in no way contradict those of Mackay (1989) and Piper (1996). As with their studies, there is very little agreement between participants filling out a similarly structured questionnaire. It is only when there is agreement between participants as to the relative importance of criteria that there is found to be similarity in their reasoning behind this view, but even then there is the issue of the relationships between constructs to consider. Real convergence in views in itself is a rare occurrence, and it can realistically be said that the findings of this study are in their own way compatible with those of the previous research. Given the lack of universal consensus in the field as a whole as to what constitutes effective teaching, it is not surprising that on a small scale such as this, teacher trainers are found not to hold identical views.

Conclusion

The conclusion was drawn that individual tutors have their own personal systems of belief with hierarchically organised ideas when it comes to judging what is effective teaching. This organisation is held together by superordinate personal constructs, that is, the principles by which people live. The principles which organise personal judgements are actually at the core of people and help them maintain their own equilibrium. Values and core principles subsume more detailed judgements. Their role should therefore be respected in helping individuals maintain their 'ontological security', trainers and trainees alike.

The implications of these findings for our team were that assessment procedures within the CertTESOL courses should take account of discrepancies between different assessors. It seems particularly important that trainees should be ensured a variety of observers during their teaching practice if the fairness of the procedure is to be optimised. Trainees' grades should continue to be decided by all those who have observed them in an open discussion of their performance in the classroom. Finally, the findings of this research were shared with the team openly, as one of the primary aims of the study was for us to deepen our awareness of our own and each others' judgements of effective teaching with a view to lessening the occurrence of conflicting advice to trainees in the future.

References

- Apelgren, B.M. 2000 'Foreign language teachers' voices: personal theories and experiences of change in Teaching English as a Foreign Language in Sweden' Unpublished PhD thesis, University of Reading

- Bogdan, R. and S. Taylor 1975 *Introduction to Qualitative Research Methods* London: John Wiley
- Borg, S. 1999 'Teachers' theories in grammar teaching' *ELT Journal* 53/3:157-164
- Cabaroglu, N. 1999 'Development of student teachers' beliefs about learning and teaching in the context of a one-year postgraduate certificate of education programme in modern foreign languages' Unpublished PhD thesis, University of Reading
- Catterall, M. and P. Maclaren 1997 'Focus group data and qualitative analysis programs: coding the moving picture as well as the snapshots' *Sociological Research Online* 2/1 [online]. Available: <<http://www.socresonline.org.uk/socresonline/2/1/6.html>>
- Cook, M.A. and H.C. Richards 1972 'Dimensions of principal and supervisor ratings of teacher behavior' *Journal of Experimental Education* 41/2:11-14
- Denicolo, P. and M. Pope 1990 'Adults learning – teachers thinking' in C. Day, M. Pope and P. Denicolo (eds.) *Insights into Teachers' Thinking and Practice* London: Falmer
- Denscombe, M. 1998 *The Good Research Guide* Buckingham: Open University Press
- Lincoln, Y.S. 1992 'Sympathetic connections between qualitative methods and health research' *Qualitative Health Researcher*. 2:375-392
- Mackay, A. 1989 'Reliability of assessments of teaching performance' Unpublished MA dissertation, University of Reading
- Mason, J. 1996 *Qualitative Researching* London: Sage
- Miles, M. and A. Huberman 1994 *Qualitative Data Analysis* London: Sage
- Piper, H. 1996 'Teacher evaluation: priorities and practices in the CTEFLA' Unpublished MA dissertation, University of Reading
- Roberts, J. 1998 *Language Teacher Education* London: Arnold
- Sendan, F. and J. Roberts 1998 'Orhan: a case study in the development of a student teacher's personal theories' *Teachers and Teaching: theory and practice* 4/2:229-244
- Weir, C. and J. Roberts 1994 *Evaluation in ELT* Oxford: Blackwell

List of Appendices

Appendix 1: Questionnaire for Trinity CertTESOL tutors

Appendix 2: Interview questions

Appendix 3: All features of classroom performance which participants rated as 5

Appendix 4: Three most important features of classroom performance

Appendix 5: Matrix displaying reasoning behind selection of three most important features of classroom performance

Appendix 6: Points in common in participants' reasoning on 'Aims'

Appendix 7: Points in common in participants' reasoning on 'Confidence building'

Appendix 1: Questionnaire for Trinity CertTESOL Tutors

Name:

Date:

Section A

The following are adapted from the Trinity CertTESOL validation guidelines for course providers. Please scan all the items first, then rate them as to their importance for successful teaching of EFL to adults at Certificate level. Please circle as appropriate.

	UNIMPORTANT			VERY IMPORTANT	
a) Identifying the needs of different types of learners	1	2	3	4	5
b) Preparation of appropriate aims for a lesson and means of achieving them	1	2	3	4	5
c) Identification and development of the attitudes and motivation of the learner	1	2	3	4	5
d) Establishing and maintaining rapport	1	2	3	4	5
e) Organisation and management of the classroom to include whole-class activity, pair, group and individual work	1	2	3	4	5
f) The ability to understand and adopt different teaching methods and styles for different learner groups	1	2	3	4	5
g) Management of learners' behaviour to encourage confidence	1	2	3	4	5
h) Management of learners' behaviour to encourage experimentation	1	2	3	4	5
i) Management of learners' behaviour to encourage consideration of the needs of other learners	1	2	3	4	5
j) Understanding and developing the role of learners in contributing to their own learning programme through self-directed study and self-evaluation	1	2	3	4	5
k) The ability to balance the requirements of accuracy and fluency as aims in teaching	1	2	3	4	5
l) Methodical treatment of errors and mistakes	1	2	3	4	5
m) The ability to balance teacher-learner participation	1	2	3	4	5
n) The giving of clear instructions	1	2	3	4	5
o) The ability to use and adapt published teaching materials effectively in whole or in part	1	2	3	4	5
p) The use of teachers' and supplementary books to complement main coursebook	1	2	3	4	5
q) Use of audiovisual aids to achieve learning aims	1	2	3	4	5
r) Devising and playing simple language games	1	2	3	4	5
s) Using authentic materials as aids to learning	1	2	3	4	5
t) Using role-play and simulations	1	2	3	4	5
u) Using song	1	2	3	4	5
v) Using video	1	2	3	4	5
w) Using drama	1	2	3	4	5

Section B

Please add any other skills and attributes which you might consider *in assessing teaching practice on the Trinity CertTESOL course*. Rate any additions according to their importance for successful teaching of EFL to adults at Certificate level. Please circle as appropriate.

	UNIMPORTANT			VERY IMPORTANT	
aa)	1	2	3	4	5
bb)	1	2	3	4	5
cc)	1	2	3	4	5
dd)	1	2	3	4	5
ee)	1	2	3	4	5
ff)	1	2	3	4	5
gg)	1	2	3	4	5
hh)	1	2	3	4	5

Continue overleaf if necessary using ii), jj) etc. to label further categories.

Section C

Taking all the criteria from Sections A *and* B (ie: including those you have added), please select what you consider to be the three *most* important for successful teaching of EFL to adults at Certificate level.

Write their letters in the boxes below as appropriate.

Taking all the criteria from Sections A *and* B (ie: including those you have added), please select what you consider to be the three *least* important for successful teaching of EFL to adults at Certificate level.

Write their letters in the boxes below as appropriate.

Thank you very much for your assistance.

Appendix 2: Interview questions

Background information

1. Code interviewee
2. Tell me about the training you have had in EFL.
3. With regard to EFL, which sectors have you worked in predominantly?

Reference to questionnaire: teaching at Certificate level

1. Can we clear up ... ?
2. I am interested in why you picked these three as being the most important criteria for successful teaching of EFL. Can you tell me why?
3. I am interested in why you picked these three as being the least important criteria for successful teaching of EFL. Can you tell me why?
4. Looking at the criteria which you categorised as being very important (5), would you like to elaborate on any of them? Reasons why?

Moving beyond Certificate level

1. What, in your view, makes good language teaching?

Appendix 3: All features of classroom performance which participants rated as 5

	<i>Participant 1</i>	<i>Participant 2</i>	<i>Participant 3</i>	<i>Participant 4</i>	<i>Participant 5</i>	<i>Participant 6</i>
Identifying learners' needs		✓	✓			✓
Aims	✓	✓	✓	✓	✓	✓
Identification and development of learners' attitudes	✓		✓			
Establishing and maintaining rapport	✓	✓	✓	✓		✓
Classroom management	✓		✓	✓	✓	✓
Adopting different teaching methods				✓		
Confidence building	✓		✓		✓	✓
Encouraging experimentation			✓			
Encouraging consideration of others' needs			✓			
Encouraging self-directed study and self-evaluation			✓			
Balancing accuracy and fluency	✓					✓
Error correction	✓	✓				
Balancing teacher-student participation	✓		✓	✓	✓	✓
Instructions	✓		✓	✓	✓	✓
Using published materials	✓		✓			✓
Using supplementary published materials			✓			
Devising language games			✓			
Using authentic materials			✓		✓	✓
Using role-play			✓			✓
Presenting structures	✓					
Asking questions	✓					
Presenting vocabulary	✓					
Using communicative activities	✓					
Reading activities	✓					
Listening activities	✓					
Writing activities	✓					
Pronunciation activities	✓					
Eliciting	✓					
Trainee self-evaluation	✓					
Language awareness: grammar		✓				
Ability to implement feedback		✓				
Ability to sequence lessons		✓				
Creating purpose-built materials					✓	
Breaking down barriers					✓	
Adapting work to suit learners					✓	
Concept checking						✓
Listening and using learner knowledge						✓
Recognising difference between teaching and testing						✓

Appendix 4: Three most important features of classroom performance

	<i>Participant 1</i>	<i>Participant 2</i>	<i>Participant 3</i>	<i>Participant 4</i>	<i>Participant 5</i>	<i>Participant 6</i>
Aims	✓	✓	✓	✓	✓	✓
Classroom management	✓					
Trainee self-evaluation	✓					
Identifying learners' needs		✓				
Language awareness: grammar		✓				
Confidence building			✓		✓	✓
Encouraging experimentation			✓			
Balancing teacher-student participation				✓		
Establishing and maintaining rapport				✓		
Breaking down barriers					✓	
Recognising difference between teaching and testing						✓

Appendix 5: Matrix displaying reasoning behind selection of three most important features of classroom performance

	<i>Participant 1</i>	<i>Participant 2</i>	<i>Participant 3</i>	<i>Participant 4</i>	<i>Participant 5</i>	<i>Participant 6</i>
Aims	<ul style="list-style-type: none"> • Unplanned lessons collapse • Trainable • Basis to evaluate their lessons • Enables learning by self-evaluation 	<ul style="list-style-type: none"> • Lessons with unclear aims meander/collapse • Encourages bottom-up planning: aims come first • Habit trainees must develop 	<ul style="list-style-type: none"> • Clear aims transform teaching • Help eliminate the unnecessary • Clear direction for learners 	<ul style="list-style-type: none"> • Lessons with unclear aims unlikely to succeed • Key thing trainees have to master 	<ul style="list-style-type: none"> • Clear progress for learners • Important for all teachers • ‘Comfort blanket’ for trainees 	<ul style="list-style-type: none"> • Means of measuring achievement • Difficult but essential skill • Learners see clear aims as important
<i>Classroom management</i>	<ul style="list-style-type: none"> • Covers many aspects of teaching • Is fundamentally what happens in the lesson • Ties in with achievement of aims 					
<i>Trainee self-evaluation</i>	<ul style="list-style-type: none"> • We learn to teach through experience • Critical awareness leads to improvement: cyclical process • Continued development beyond course 					
<i>Identifying learners’ needs</i>		<ul style="list-style-type: none"> • Informs content, pitch and aims of lesson 				
<i>Language awareness: grammar</i>		<ul style="list-style-type: none"> • Grammar is fundamental to language teaching • Trainees must understand what they are teaching • Poor knowledge can destroy learners’ trust in teacher 				
<i>Confidence building</i>			<ul style="list-style-type: none"> • Self-confidence leads to experimentation • Willingness to accept mistakes • Confidence is empowering 		<ul style="list-style-type: none"> • Learners perform best orally when confident • Willingness to experiment • Risk taking leads to increased knowledge of language 	<ul style="list-style-type: none"> • Confident students give and therefore learn • Supportive learning environment means students learn from mistakes • Lack of confidence is a barrier to learning

...Continued

	Participant 1	Participant 2	Participant 3	Participant 4	Participant 5	Participant 6
Encouraging experimentation			<ul style="list-style-type: none"> • Ties in with confidence • Mistakes lead to learning 			
Balancing teacher-student participation				<ul style="list-style-type: none"> • Monopolising classroom time is common trainee error 		
Establishing and maintaining rapport				<ul style="list-style-type: none"> • Lack of rapport leads to lack of progress • Fundamental to being a good teacher • Learners need a sense of teacher's individuality 		
Breaking down barriers					<ul style="list-style-type: none"> • Barriers between teacher and learners are destructive • Teacher can maintain respect as part of the group • Teacher is not the font of all knowledge • Learners want teachers to be facilitators of learning 	
Recognising difference between teaching and testing						<ul style="list-style-type: none"> • Fundamental shift trainees have to make • Runs through many areas in teaching • One of the keys to success

Appendix 6: Points in common in participants’ reasoning on ‘Aims’

Results of having unclear aims	Participant 1 Participant 2 Participant 4	<ul style="list-style-type: none"> • Unplanned lessons collapse • Lessons with unclear aims meander/collapse • Lessons with unclear aims unlikely to succeed
Skill to be learned	Participant 1 Participant 2 Participant 4 Participant 5 Participant 6	<ul style="list-style-type: none"> • Trainable • Habit trainees must develop • Key thing trainees have to master • Important for all teachers • Difficult but essential skill
Providing direction	Participant 3 Participant 5	<ul style="list-style-type: none"> • Clear direction for learners • Clear progress for learners
Evaluation	Participant 1 Participant 6	<ul style="list-style-type: none"> • Basis to evaluate their lessons • Enables learning by self-evaluation • Means of measuring achievement

Appendix 7: Points in common in participants’ reasoning on ‘Confidence building’

Experimentation and learning	Participant 3 Participant 5	<ul style="list-style-type: none"> • Self-confidence leads to experimentation • Willingness to experiment • Risk taking leads to increased knowledge of language
Acceptance of mistakes	Participant 3 Participant 6	<ul style="list-style-type: none"> • Willingness to accept mistakes • Supportive learning environment means students learn from mistakes
Benefits to learner	Participant 3 Participant 5 Participant 6	<ul style="list-style-type: none"> • Confidence is empowering • Learners perform best orally when confident • Confident students give and therefore learn • Lack of confidence is a barrier to learning