

A content analysis method to measure critical thinking in face-to-face and computer supported group learning

D. R. Newman, Brian Webb and Clive Cochrane
Queen's University Belfast, Information Management Dept.

Abstract

This paper gives a detailed account of the content analysis method developed at Queen's University Belfast to measure critical thinking during group learning, as used in our controlled comparisons between learning in face-to-face and computer conference seminars. From Garrison's 5 stages of critical thinking, and Henri's cognitive skills needed in CMC, we have developed two research instruments: a student questionnaire and this content analysis method. The content analysis relies on identifying, within transcripts, examples of indicators of obviously critical and obviously uncritical thinking, from which several critical thinking ratios can be calculated.

Introduction

How can we evaluate Computer Supported Co-operative Work (CSCW), in particular that kind of work we call learning? This paper presents a content analysis method we have used in evaluating the quality of group learning taking place in seminars held both face-to-face and with computer mediated communication (CMC).

It is part of a series of papers (Webb, Newman and Cochrane 1994; Newman 1994) resulting from our controlled experiments, where half of the seminars in an Information Society module were held face-to-face, the other half over the Network Telepathy asynchronous computer conferencing system. Each seminar group used both media to discuss issues raised in the course lectures and readings. The key educational objective of these conventional seminars was to encourage critical thinking among the students about controversial issues in IT and society, since these have no one right answer that the student can memorise for the examinations.

We have found evidence for critical thinking (as explained below) in both face-to-face and computer conference seminars, using both a student perception questionnaire (Webb, Newman and Cochrane, 1993) and the content analysis technique described in this paper. The content analysis revealed similar depths of critical thinking on several different indicators, although we found more new ideas emerged in face-to-face seminars, and more ideas in the computer conferences were important, justified or linked together.

This paper elaborates on the content analysis method we used to measure critical thinking in CSCL, in sufficient detail to enable others to replicate (and improve upon) it. It should be applicable in any learning situation where a transcript of group learning discussions can be obtained (not just in conventional seminars), and where the object is to evaluate the learning process taking place, rather than assessing individual student performance.

Current approaches to evaluating CSCL

In September 1993 there was a workshop on evaluating Computer Supported Co-operative Work, held at the start of the European conference on CSCW. Some of the presentations amounted to stories or descriptions, picking out striking statements or events from a particular application. Some others tried to measure everything they could easily measure, then came to the

meeting hoping for suggestions as to what they should do with the data. Not one groupware developer turned up. Similarly, Robin Mason's review of evaluation methodologies for computer conferencing applications found that most of the literature is made up of descriptions of applications; and that "most computer conferencing research stops with quantitative analyses of messages sent and by whom, number and time of logons, and message maps showing the number of replies and message chains" (Mason 1991). Theories of what constitutes good work, or even control groups, are uncommon.

More considered evaluation approaches tend to follow different disciplinary perspectives, such as:

Sociology of working life: How society and CSCW-using organisations interact, and how different power groups within these organisations control, benefit and suffer from the computer tools. (See, e.g., Wagner 1993). This approach is used in both the socio-technical and human-centred systems design traditions (Nurminen 1988), and has been taken up by trade unions (URCOT 1993).

Ethnomethodology: using anthropological techniques, people have carried out detailed studies of CSCW in particular organisations, avoiding, at all costs, imposing external theories when analysing the hours of videotape (When Lucy Suchman gave a presentation at Lancaster University on her work at an airport ground control office, she never once mentioned the word power.) These give insights into human-human and human-computer interactions, the dynamics of social processes, and breakdowns in the system. (Suchman 1987; Bowers and Benford 1991)

Network and workflow analyses: of communications or information flows within organisations. By creating network diagrams showing the amount of communication between different people it is possible to identify problems with human and computer supported systems (such as people isolated from group decision-making). But this says nothing about the quality of the work done.

HCI design: Human-Computer Interaction-driven analyses of CSCW implementation to give guidance to designers (Urquijo 1993, Macleod 1993). These concentrate on computer interaction rather than human-human interaction, and often on individual task analysis rather than group collaboration. Questions about whether it is worth using the system at all are outside the scope of such analyses.

Cognitive and Social Psychology: How CSCW affects the quality of work done in groups, individual working styles, and motivation for and against CSCW. It is this approach that we have adopted, since our prime concern is whether the quality of work (in our case learning) can be maintained through using CSCL in the face of increasing class sizes.

Reviews of the literature on the evaluation of CSCW, and especially CSCL have been published by Mason (1991) and by Collis (1994). Collis notes a shift in the focus of co-operative learning research to emphasise theoretically grounded instrumentation and methodology.

Mason found that surveys, user interviews, empirical experimentation, case studies and computer generated statistical measurements are being used to evaluate computer conferencing. None of these tell us much about the quality of student learning taking place. She suggests widening the net "by using other educational goals, such as collaborative learning, critical thinking, deep understanding of course material or broad awareness of issues, and by breaking these down into examples of behaviour and written work which display these characteristics, it is possible to analyse conference content and draw conclusions about the educational value of the particular on-line activity" and concludes with a plea "for evaluators to take up the challenge of content analysis both as a key to increasing the professionalism of the field and as the essence of the educational value of the activity".

We have taken up this challenge to develop research instruments based on theories of:

Group learning, deep learning and critical thinking

From educational research we know that learners can adopt deep or surface learning approaches. Surface learning approaches include skimming, memorising, and regurgitating for tests, whereas deep learning requires a critical understanding of material. Deep learning is promoted by active learner participation. Biggs (1985) associated deep learning approaches with 'affective involvement' which is supported by interaction.

This is in line with other work that sees student learning as an individual but as a social phenomena. Psychologists have long believed that individual cognitive skills are developed in a social context (Resnick et al 1991) and have cited the psychological development of children as evidence (Rogoff 1990). The importance of the social context to learning has been emphasised by Lipman (1991) who believes that the development of a 'community of enquiry' is essential for the development of higher level, critical thinking skills within the individual. Entwistle and Ramsden (1983) have highlighted the importance of horizontal (student-student) and vertical (student-teacher) interaction.

A clear link between critical thinking, social interaction and deep learning has emerged.

Group learning is a good way of encouraging such social interaction, and has often been used to promote deep learning. To this end, best educational practice provides techniques for getting group work in large classes (e.g. rounds, line-ups, pyramids, projects, courts of enquiry, posters, brainstorming), critical peer and self-assessment, resource-based individual and group learning--without a computer in sight.

A good CSCL system should do as well as such face-to-face group learning in promoting the use of deep learning styles. The problem is: how do we evaluate the quality of the learning activities going on in both situations?

We could try to measure deep learning. However it is not clear that everyone is using deep learning to refer to the same concept or behaviour. Are deep and surface learning extremes of individual learning styles, or alternative strategies which a student adopts according to the needs of the moment? There have been a number of scales designed to measure deep learning. However, they have not proved robust when tested outside their original context, culture or educational level. Richardson (1993) has found that none of the three different student learning inventories "measures what it claims in the sense that its constituent structure can be empirically confirmed on participants similar to those on whom it was originally developed". Nor have they been specifically designed to assess group learning in a social context. So we looked at the key link between group learning and deep learning--critical thinking. Successful group problem-solving processes require critical thinking, leading to the critical understanding needed for deep learning.

Dimensions of CSCL evaluation

Henri (1991) has identified 5 dimensions on which we can evaluate CMC.

1. *participative*: at a minimum, is it used? This can degenerate to the usual bean counting of numbers and lengths of messages, so despised by Mason. It is easy to measure (at least in CMC, it is harder face-to-face), but says nothing about the quality of what is going on.
2. *social*: this is clearly important, since the social dimension provides some of the motivation for people to make use of the system. And in certain circumstances this can be the prime goal of the system, such as when the participants are communicating between the Falls Road and the Shankill. However, this dimension says nothing about the quality of learning taking place, so it is outwith the main purpose of the discussion
3. *interactive*: it is possible to measure the interactions, responses and commentaries taking place, seeing in detail how particular events or statements lead to particular responses. Such analyses, by whatever techniques (scripts, natural language analysis, ethnography) can give useful insights into how to improve conversations (through facilitation, for example), but do not tell us much about the type of learning that has gone on.
4. *cognitive*: people use skills connected to reasoning which uses critical thought. It is this dimension which is of most interest to us as educators, since in many of our courses at QUB we wish to encourage critical thinking, rather than shallow search for 'the one right answer'.

5. *metacognitive*: reasoning about reasoning and self-awareness. Although Henri distinguishes this from the cognitive dimension, I would not consider this an important distinction. Artificial Intelligence makes the same distinction between knowledge and meta-knowledge. But as soon as you start to implement meta-knowledge (e.g. problem-solving strategies) in an expert system, you find you are just adding more rules, or extra levels in classification hierarchies. It becomes more useful to think of abstraction levels, of applying cognitive skills to more abstract chunks of knowledge than a simple binary declarative/procedural distinction. At what point does criticising an approach to criticising someone's criticism become metacognitive?

In our work, we set out to find good ways of evaluating the cognitive dimension, the one most relevant to the learning activities and styles we wished to encourage, namely:

Evaluating critical thinking.

We need to look for signs of critical thinking in a social context. This is quite different from measuring student performance, student participation or system usability. Nor is it as narrow as Ennis' concept of critical thinking as the correct assessing of statements (Ennis 1967), since that is only one stage that individuals go through in group critical thinking processes. Critical thinking is not just limited to the one-off assessment of a statement for its correctness, but a dynamic activity, in which critical perspectives on a problem develop through both individual analysis and social interaction. Garrison (1992) emphasised this, in saying, "Thus, it is in the shared world that true meaning is achieved. While constructing meaning is a personal responsibility, the process of critical thinking also includes the application of meaning structures to the specifics of the context. This is, if meaning is to be more than belief it must go beyond simply internal reflection. The truth of concepts is determined through collaborative action which necessitates sharing control of the process."

In the fields of Operations Research and Cognitive Science there are models of such dynamic processes, such as the Intelligence, Design, Choice and Implementation model of decision-making famous in the Decision Support System literature. Brookfield (1987) suggested a five-phase model of critical thinking in adult education: a triggering event, situation appraisal, exploration to explain anomalies, development of alternative perspectives, and integrating perspectives into the fabric of living.

Most relevant to the evaluation of critical thinking in CSCL is Garrison's (1992) model of critical thinking as a 5-stage process. For the most part Garrison's stages correspond closely to the cognitive skills Henri recognises as important to the cognitive dimension of CMC. Garrison considers the critical thinking process as a problem-solving one. The stages are:

Stage 1. Problem identification Skill 1. Elementary clarification

Based on an initial motivation to learn, a 'triggering event' arouses and sustains interest and curiosity stimulated by interaction with others. Learners observe or study a problem, identify its elements, and observe their linkages in order to come to a basic understanding.

Stage 2. Problem definition Skill 2: In-depth clarification

Framing the problem and an approach to its solution using the experience of others. Learners analyse a problem to come to an understanding which sheds light on the values, beliefs and assumptions which underlie the statement of the problem.

Stage 3. Problem exploration Skill 3. Inference

Getting insights and understanding based on self and group learning. The skills needed to extend beyond the basis definition include inference: induction and deduction, admitting or proposing an idea on the basis of its link with propositions already admitted as true. But they also include the creative skills needed to widen the field of possible solutions.

Stage 4. Problem evaluation/applicability Skill 4. Judgement

The evaluation of alternative solutions and new ideas within a social context. This needs judgemental skills of making decisions, statements, appreciations, evaluations and criticisms or "sizing up".

Stage 5. Problem integration Skill 5. Strategy formation

Proposing co-ordinated actions for the application of a solution, or following through on a choice or decision. This draws upon existing personal knowledge but is then validated within the group. This is the stage where the solutions are grounded back in the real world.

Given such a model of critical thinking, it is possible to design methods to evaluate it. We have used two methods: student questionnaires and content analysis. In Webb, Newman and Cochrane (1994) we described how we designed and administered a student perception questionnaire designed to measure:

1. The capacity of the computer conferencing environment to arouse and sustain interest and to increase awareness of important issues.
2. Problem definition. To what extent does computer conferencing clarify course objectives and make relevant personal experience?
3. Problem exploration. Does computer conferencing have the capacity to help the student to develop new ideas and solutions, understand issues including those contained within study texts?
4. Critical assessment of course content, possible solutions and the ideas of others.
5. Problem integration. To what extent does computer conferencing help the student to understand by applying course content to his or her own life situation?

We found evidence for critical thinking in both face-to-face and computer conference seminars, and some confirmation for Garrison's model through factor analysis of the questionnaires.

The questionnaire assessed student perceptions over a semester. To get a more detailed understanding of the quality and type of learning taking place, you can use content analysis. For that we needed to identify textual indicators of critical thinking.

Critical thinking indicators

Mason (1991) proposed that the analysis of conference messages should be based on the educational value they exhibit. In particular:

- do the participants build on previous messages?
- do they draw on their own experience?
- do they refer to course material?
- do they refer to relevant material outside the course?
- do they initiate new ideas for discussion?

These are all indicators of critical thinking (at different stages). Henri (1991) identified indicators of each of her cognitive reasoning skills. For example, her indicators of judgement skills were: judging the relevance of solutions, making value judgements and judging inferences. In principle it should be possible to classify statements in a discussion transcript by Henri's cognitive skills (or Garrison's stages) according to her indicators. Henri encountered difficulties, "superficial results telling us only the presence and frequency of use of these skills". I would suggest that this is due to three problems:

1. Individuals in a group discussion are often at different stages in Garrison's process of critical thinking. So it is hard to trace a consistent movement through the stages through computer conference content analysis, although it might be possible in other groupware systems.
2. Henri's indicators of the cognitive skills are fairly broad, and can include a number of different activities. The well known examiner's (and expert system) trick, of dividing subtle value judgements into several simpler well-defined criteria, can be applied to them.
3. The indicators do not attempt to evaluate the depth of these cognitive skills, e.g. to distinguish between critical value judgements and uncritical statements of values.

Henri's solution to (3) brings us back to deep and surface learning (or in-depth and surface processing). She gives a list of paired opposites, one an indicator of surface processing, one of in-depth processing. For example, 'making judgements without offering justification', versus 'setting out the advantages and disadvantages of a situation or solution'. On examining that list, you find these are indicators of critical versus uncritical thinking, at different stages of the critical thinking process.

We developed our own set of paired indicators, by simplifying Henri's pairs (as in b, above), by looking for indicators in all of Garrison's stages (including opposites derived from our and Anderson's (1993) questionnaires), and from our experience of using similar techniques for assessing student work in computer conferences. These are shown in Fig. 1.

R+- Relevance

- R+ relevant statements
- R- irrelevant statements, diversions

I+- Importance

- I+ Important points/issues
- I- unimportant, trivial points/issues

N+- Novelty. New info, ideas, solutions

- NP+ New problem-related information
- NP- Repeating what has been said
- NI+ New ideas for discussion
- NI- False or trivial leads
- NS+ New solutions to problems
- NS- Accepting first offered solution
- NQ- Squashing, putting down new ideas
- NQ+ Welcoming new ideas

NL+

learner (student) brings new things in

NL-

dragged in by tutor

O+- Bringing outside knowledge/experience to bear on problem

OE+

Drawing on personal experience

OC+

Refer to course material

OM+

Use relevant outside material

OK+

Evidence of using previous knowledge

OP+

Course related problems brought in (e.g. students identify problems from lectures and texts)

OQ+

Welcoming outside knowledge

OQ-

Squashing attempts to bring in outside knowledge

O-

Sticking to prejudice or assumptions

A+- Ambiguities: clarified or confused

AC+

Clear, unambiguous statements

AC-

Confused statements

A+

Discuss ambiguities to clear them up

A-

Continue to ignore ambiguities

L+- Linking ideas, interpretation

L+

Linking facts, ideas and notions

L+

Generating new data from information collected

L-

Repeating information without making inferences or offering an interpretation.

L-

Stating that one shares the ideas or opinions stated, without taking these further or adding any personal comments.

J+- Justification

JP+

Providing proof or examples

JS+

Justifying solutions or judgements

JS+

Setting out advantages and disadvantages of situation or solution

JP-

Irrelevant or obscuring questions or examples

JS-

Offering judgements or solutions without explanations or justification

JS-

Offering several solutions without suggesting which is the most appropriate.

C+- Critical assessment

C+

Critical assessment/evaluation of own or others' contributions

C-

Uncritical acceptance or unreasoned rejection

CT+

Tutor prompts for critical evaluation

CT-

Tutor uncritically accepts

P+- Practical utility (grounding)

P+

relate possible solutions to familiar situations

P+

discuss practical utility of new ideas

P-

discuss in a vacuum (treat as if on Mars)

P-

suggest impractical solutions

W+- Width of understanding (complete picture)

W-

Narrow discussion. (Address bits or fragments of situation. Suggest glib, partial, interventions)

W+

Widen discussion (problem within a larger perspective. Intervention strategies within a wider framework.)

(+- is an attempt to render the plus-or-minus sign in ASCII)

Fig. 1 Indicators of critical (+) and uncritical (-) thinking

Given the problem of identifying Garrison's stages of critical thinking in a group discussion when different participants are at different stages, these indicators have not been grouped by critical thinking stage, but by common behaviours.

The indicators have been chosen as fairly obvious opposites. It should not be difficult for an evaluator to identify statements illustrating these extremes.

Analytical method

We prepared transcripts of tape-recorded seminars, and automatically stored transcripts of computer conferences, as part of a controlled experiment in which student groups carried out half their seminars using each technique (Webb, Newman and Cochrane 1994, Newman 1994).

Rather than classify every statement in a transcript as, e.g. critical assessment or uncritical acceptance, we mark and count the obvious examples, and ignore the intermediate shades of grey. This eases the task of the assessors, since there is less need for subtle, subjective, borderline judgements.

These statements may be phrases, sentences, paragraphs or messages containing one unit of meaning, illustrating at least one of the indicators. Of course, one statement might show more than one indicator, such presenting a new idea which widens the discussion. Or indicators can even overlap. Fig. 2 and 3 show example fragments of marked-up transcripts.

```
=====  
infosociety/monday1 #6, from sandrap, 479 chars, Fri, 30 Apr 93  
10:57  
-----  
<C+<I believe that through the computer an affiliation  
between  
the workers can occur and users could have a team spirit. The  
question is wether this type of interaction is more beneficial  
than face to face or would an office based team be more  
successful?>C+><R+I+OC+<This continues my opinion  
from Mondays  
seminar that a joint office and teleworking based operation would  
be better. <JS+<Here you have the advantages of meeting  
face to  
face as a team and working on the system together as a  
team.>JS+>>R+I+OC+>
```

Fig. 2 Example of critical thinking indicators in a computer conference transcript

```
<OM+C+JS+<patrick: the, the thing about it is, is the  
exchange is  
underneath the barracks, it's actually in the barracks, because  
the exchange it's the only place they can keep it safe. So that's  
why people get very...>OM+C+JS+>
```

```
<C+<andrea: Yeah, but that's over here, that's  
different.>C+>
```

```
<C+<patrick: Yeah, but why, why should we not care? I mean  
it's  
only when it's Britain itself that people complain...>C+>
```

```
<C+OM+<andrea: Because there's dickheads running about out  
there
```

who like blowing up telephone exchanges...>C+OM+>

<C-<patrick: Oh I know that, I know that...>C->

Fig. 3 Example of critical thinking indicators in a seminar transcript

Once the scripts are marked, the totals for each + or - indicator are counted, and a critical thinking ratio calculated for each, x ratio = $(x+ - x-)/(x+ + x-)$, converting the counts to a -1 (all uncritical, all surface) to +1 (all critical, all deep) scale. This was done to produce a measure that was independent of the quantity of participation, reflecting only the quality of the messages. Fig. 4 shows how this works for a summary analysis of our first results.

Fig. 4 Calculating critical thinking ratios

Scoring criteria	Ratios		CC		f2f	
	CC	f2f	+	-	+	-
R+- Relevance	0.88	0.85	48	3	137	11
I+- Importance	0.89	0.35	33	2	29	14
N+- Novelty. New info, ideas, solutions	0.37	0.59	24	11	65	17
A+- Ambiguity and clarity/confusion	*	-0.05	2	2	20	22
O+- Bringing outside knowledge/ experience to bear on problem	1.00	0.76	35	0	89	12
L+- Linking ideas, interpretation	0.80	0.06	18	2	27	24
J+- Justification	0.69	0.48	27	5	80	28
C+- Critical assessment	0.89	0.93	35	2	143	5
P+- Practical utility (grounding)	*	0.88	1	0	15	1
W+- Width of understanding	*	0.06	1	0	9	8

CC=computer conferencing. f2f=face-to-face. * Not calculated because sample is too small.

Results of the analysis

A detailed analysis of our results will appear in a later paper. Some brief points are included here as an example of the usefulness of this content analysis technique. See Fig. 5, in the Appendix, for a spider diagram of these ratios.

Initial analysis of our first semester's experiment showed that these ratios for our face-to-face and computer conferencing seminars, were similar, showing critical thinking in both. There were notably more positive ratios in computer conferences for important statements and linking ideas, and slightly less for novelty.

A possible explanation for the lack of new ideas is that the asynchronous computer conferencing environment discouraged students from contributing novel, creative ideas (as in brainstorming), but rather encouraged considered, thought out contributions. The slightly more +ve ratios for outside knowledge and justification are consistent with this: a statement of opinion in a face-to-face discussion becomes an evidentially justified point in a computer conference message.

Against this, the Telepathy computer conferencing system made it easy for the students to look though several previous messages before making a comment, and then link together different ideas. In a face-to-face seminar, students would have to remember earlier points after the discussion has moved on.

The Telepathy computer conferencing environment appears to have hindered the earlier exploratory and creative stages of Garrison's problem-solving critical thinking process, but helped the later evaluation and integration stages. This has clear educational implications. In my case, it has led to a search for different face-to-face and computer-supported techniques to support creative ideas generation, while explicitly setting evaluation and integration tasks to be carried out using the computer conferencing system. When implemented, these will be evaluated using the same methods.

Practical content analysis problems

This content analysis technique works well for the transcripts we have tried it on. But we did find some practical problems.

Since some of the indicators, such as OC+, OM+, I+, R+ rely on subject knowledge, the scripts need to be marked by someone with such knowledge: ideally the class tutor. So getting multiple evaluators to control for subjective scoring is not usually possible. Instead, just as in an expert system, one relies on the experience and expertise of the tutor. As others replicate this work, it should be possible to get a clearer idea of what variability this could introduce. Where similar statements are scored differently, presenting them to different experts for a repertory grid analysis should clarify our constructs, and suggest better pairs of indicators on which to measure critical thinking.

In our transcripts, we found few obvious A+-, P+- or W+- statements. This may be an effect of the nature of the student task: general discussion rather than the more specific problem-solving activities envisaged by Garrison. But if others find the same in other learning situations, these indicators should be dropped.

Some teachers find picking out examples of uncritical thinking hard; particularly those who assess work only by looking for positive points. Ideally new scorers should practice blind marking of already scored transcripts before starting on their own class. This content analysis technique is a skill that evaluators need to develop, rather like marking examinations.

Working out where statements or points start and end can be hard in face-to-face transcripts, as people interrupt each other, continue across interruptions and so on. It is rarely a problem in CSCL transcripts.

At present, content analysis is long-winded. Not as much as transcript analysis in knowledge acquisition, where it takes 8 hours to analyse one hours interview, but it still takes too long for this to be used outside research projects. We have attempted to reduce the time needed (compared to Henri) by closely focusing on cognitive factors, in particular, critical thinking, and by attempting to identify fairly obvious indicators. But a faster process is needed before content analysis of critical thinking could be used in ordinary classroom evaluation.

Conclusion

Henri was concerned to make the content analysis procedure more scientific, resistant to subjective manipulations by evaluators and/or teachers, based on rigorous analysis of evaluation aims and on a strong theoretical framework.

Our aims were to measure the amount and type of critical thinking taking place in group learning, in order to check on the possibilities of using CSCL to promote deep learning even in large classes. We did not attempt to measure the interaction taking place, the amount of participation or the social dimension. Once we can measure the quality of learning, then we can worry about how this is affected by social or interactive factors. The content analysis methodology described here meets our aims pretty well, in producing -1 to +1 scales of different aspects of critical thinking.

Garrison's theory of critical thinking provides a good theoretical framework for analysing critical thinking in a group learning context. Fairly directly it can be used to develop student questionnaires. And it provides a guide, together with Henri's work, from which a set of indicators of critical thinking can be identified. From these we can assess how much and what kind of critical thinking is taking place.

Both techniques have some element of subjectivity, whether of students filling in questionnaires or evaluators scoring transcripts. But we have gone some way to reduce (not eliminate) this in our content analysis, through the choice of indicators, through our analytical method, and by the use of control groups. We invite others to use these techniques, and improve upon them.

References

- Anderson T. (1993)
Socially shared cognition in distance learning: An exploration of learning in an audio teleconferencing context. Unpublished doctoral dissertation proposal, University of Calgary.
- Bevan, N. and M. Macleod (1993)
Usability Assessment and Measurement. In: M. Kelly (Ed.) *The Management and Measurement of Software Quality*, Ashgate Technical/Gower Press.
- Biggs, J. B. (1987)
Student approaches to studying and learning. Melbourne: Australian Council for Educational Research.
- Bowers, John and S. Benford (eds) (1991)
Studies in Computer Supported Cooperative Work: Theory, Practice and Design. Amsterdam: Elsevier.
- Brookfield, S. D. (1987)
Developing critical thinkers. San Francisco: Jossey-Bass.
- Collis, Betty A. (1994)
Collaborative learning and CSCW: research perspectives for internetworked educational environments. in R. Lewis and P. Mendelsohn (eds.) *Lessons from Learning*, Elsevier Science.
- Corbett, M, M. Macleod and M. Kelly (1993)
Quantitative Usability Evaluation - the ESPRIT MUSiC Project, Proc. of HCI International '93 Conference, Florida, USA, Aug. 1993.
- Ennis, R. H. (1967)
A concept of critical thinking. pp. 114-48 in: B. Paul Komisar and C. J. B. Macmillan (eds) *Psychological concepts in education*. Chicago: Rand McNally.
- Entwhistle. N and Ramsden P. (1983)
Understanding Student Learning. London: Croom Helm.
- Garrison, D. R. (1992)
Critical thinking and self-directed learning in adult education: an analysis of responsibility and control issues. *Adult Education Quarterly*, 42(3) 136-148
- Henri, F. (1991)
Computer conferencing and content analysis in O'Malley, C. (ed.) *Computer Supported Collaborative Learning*: Heidelberg: Springer-Verlag.
- Lipman, M. (1991)
Thinking in Education. Cambridge: Cambridge University Press
- Macleod, Miles (1993)
Position paper: evaluation methods for collaborative work, ECSCW evaluation workshop, Milan, 1993, from . For detailed accounts of the MUSiC methodology, see Bevan 1993, Corbett 1993, Maissel 1993 and Renger 1992.
- Maissel, J, M Macleod et al. (1993)
Context Guidelines Handbook 2.1, National Physical Laboratory, Teddington, UK.
- Mason, R. (1991)
Methodologies for evaluating applications of computer conferencing in Kaye, A.R.(ed) *Collaborative Learning through Computer Conferencing*: Heidelberg: Springer-Verlag.
- Newman, D. R. (1994)
Evaluating the quality of learning in CSCL. Presented at UK CSCW SIG workshop on CSCL, Heriott-Watt University, Edinburgh, March 1994.
- Nurminen, Markku I. (1988)

People or computers: three ways of looking at Information Systems. Lund: Studentlitteratur/ Chartwell-Bratt.

Renger, R, M Macleod et al. (1992)

MUSiC Performance Measurement Handbook. National Physical Laboratory, DITC, Teddington, UK.

Resnick, L., Levine, J. & Teasley S. (eds) (1991)

Perspectives on socially shared cognition (pp. 384-397). Washington: American Psychological Assoc.

Richardson, John T. E. (1994)

Using questionnaires to evaluate student learning: Some health warnings, pp. 73-88 in: Graham Gibbs (ed.) Improving Student Learning: Theory and Practice, Oxford Centre for Staff Development, Oxford Brookes University.

Rogoff, B. (1990)

Apprenticeship in thinking: cognitive developments in social context, New York: Oxford University Press.

Suchman, Lucy (1987)

Plans and situated actions: the problem of human-machine communication. Cambridge: Cambridge University Press.

URCOT (1993)

Huron etc. Project Briefing Working Paper No. 3, Union Research Centre on Office Technology, Melbourne.

Urquijo, Silvia Ponguta, Stephen A. R. Scrivener and Hillary K. Palmen (1993)

The use of breakdown analysis in synchronous CSCW system design, ECSCW, Milan, Sept. 1993.

Wagner, Ina (1993)

Fragmented exchange: Why should complex organisations implement CSCW technology?, ECSCW Evaluation Workshop position paper, from .

Brian Webb, D. R. Newman and Clive Cochrane (1994)

Towards a methodology for evaluating the quality of student learning in a computer-mediated-conferencing environment in Graham Gibbs (ed.) Improving Student Learning: Theory and Practice. Oxford: Oxford Centre for Staff Development, Oxford Brookes University. 1st International Symposium Improving Student Learning: Theory and Practice, Warwick University, Sept. 1993.

Appendix

Content analysis of Information Society online and face-to-face seminars, by indicators of deep vs. surface learning.

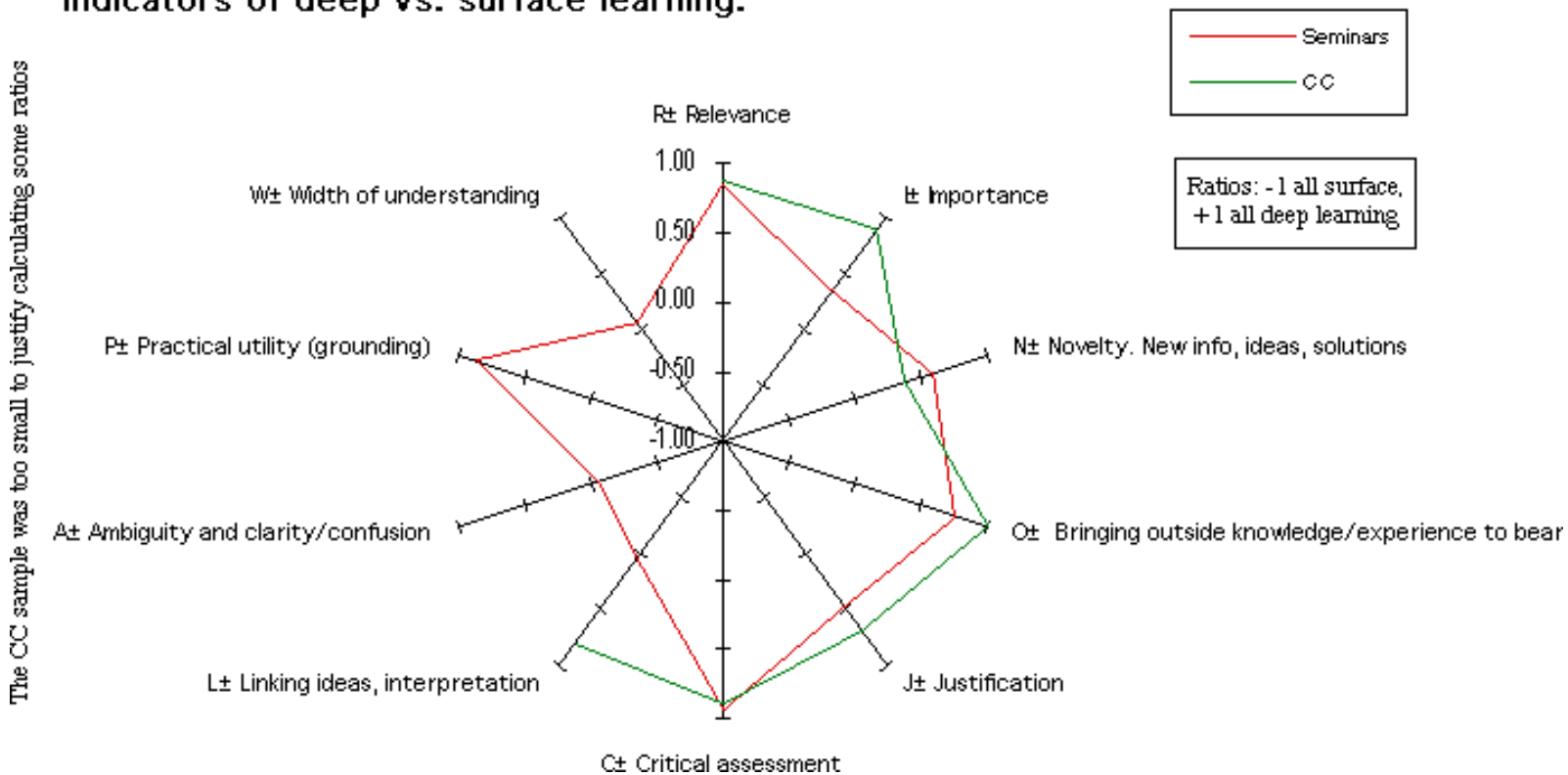


Fig. 5 Comparing critical thinking ratios for face-to-face seminars and computer conferences.