DOCUMENT RESUME

ED 388 270 IR 017 419

AUTHOR	Major, Nigel
TITLE	Evaluating COCA--What Do Teachers Think?
PUB DATE	94
NOTE	7p.; In: Educational Multimedia and Hypermedia, 1994. Proceedings of ED-MEDIA 94--World Conference on Educational Multimedia and Hypermedia (Vancouver, British Columbia, Canada, June 25-30, 1994); see IR 017 359.
PUB TYPE	Reports - Descriptive (141) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Artificial Intelligence; Authoring Aids (Programming); Case Studies; *Computer Assisted Instruction; *Educational Strategies; Elementary Secondary Education; Foreign Countries; Higher Education; Instructional Effectiveness; *Instructional Material Evaluation; *Intelligent Tutoring Systems; Questionnaires; *Teacher Attitudes
IDENTIFIERS	*COCA (Authoring Aid)

ABSTRACT

COCA, which consists of both authoring tools and a runtime shell, is a system intended to provide teachers with genuine access to intelligent tutoring system (ITS) technology and to give them control over domain material and teaching strategies. To evaluate the effectiveness of COCA, 10 subjects (five university teachers and five school teachers) were given an authoring task and asked to complete a questionnaire. The subjects were presented with domain material (the American Revolution), an initial teaching strategy and a meta-teaching strategy in the form of prepared knowledge bases, and a set of instructions for altering the teaching behavior of the final system and extending the domain material. The questionnaire covered the general use of COCA, the suitability of the teaching strategy representation, the domain representation, and how subjects' opinions of AI (artificial intelligence) in general and COCA in particular had changed as a result of the task. Results indicate that school teachers' attitudes toward AI improved and university teachers' attitudes remained positive; COCA was found to be successful for simple tutoring systems, yet too difficult to use. (Contains 17 references.) (AEF)

Evaluating COCA - What do teachers think?

Nigel MAJOR
Department of Psychology, University Park, Nottingham. NG7 2RD. UK
Email: nigel@psyc.nott.ac.uk

Abstract: This paper discusses an evaluation of COCA, a system that gives teachers control over domain material, teaching strategy and meta-teaching strategy. The purpose of this evaluation is to study more fully the effectiveness of the system. Ten subjects were given an authoring task. The resulting knowledge bases, together with a questionnaire, made up the experimental results. This study shows the strengths and weaknesses of the COCA approach and whether it has helped improve teachers' attitudes towards AI. The results show that the system was successful, yet too complex. Results are tentative due to the size of the experiment.

1 Introduction

A lot of work has investigated the efficacy of different teaching styles in the classroom. Examples include the Plowden Report (1967), looking at teaching methods used in primary schools, studies of formal and informal teaching styles, such as Bennett (1976), and studies of the teaching styles of computer-based training (Eaton & Olson, 1986). This work suggests that many teachers have rejected educational software as it uses the wrong teaching style and is not available in the right subject area. Only recently have intelligent tutoring systems (ITSs) really been able to address these problems, as systems have begun to emerge which offer a range of teaching styles (Spensley, Elsom-Cook, Byerley, Brooks, Federici, & Scaroni, 1990) with tools allowing the teacher to change those styles. Consequently, there is little or no actual evaluation of ITSs in the classroom which support the teacher with a range of teaching strategies, and few or no empirical studies have been published (Murray, 1993). The issue of evaluation has been highlighted as needing attention for ITS systems in general (Brna, Ohlsson & Pain, 1993), with an increasing need for researchers to use appropriate methods for evaluating their results.

COCA, which consists of both authoring tools and a runtime shell, is a system intended to provide teachers with genuine access to ITS technology. Teachers are either uninformed about using AI techniques in schools, or do not have the resources to use them, or more seriously, do not trust the decision making of a system which would be controlling some of the teaching in their classroom. As the result of a number of interviews and lesson observations, Major (1993a) reported requirements from teachers about what they need from an intelligent assistant. Predominant amongst these was the ability to control the teaching style of the system. As a result of these requirements, and also on the basis of the use of meta-level reasoning in the fields of planning and knowledge-based systems research, a prototype system COCA-0, was built. This was then informally evaluated with teachers, resulting in a more complete system, COCA-1, described in Major & Reichgelt (1992), which gives a great deal of control over the teaching style of the system to the teachers themselves. The initial evaluation of COCA was the reconstruction of the teaching strategies of other systems described in Major (1993b). Further, in Major & Reichgelt (1991) a teacher used COCA-1 to build a simple ITS for algebraic equations. This case study demonstrated that a teacher can develop a simple ITS with COCA's authoring tools.

This paper describes a more formal evaluation of COCA-1, comparing school teachers and university teachers, and contrasting their attitudes towards using AI in the classroom before and after the experiment. The subjects in the experiment were asked to perform a given authoring task and to fill in a questionnaire. Both the knowledge bases produced during the task, and the completed questionnaires, were analysed against the aims of the experiment. The purpose of the experiment was to evaluate a more detailed use of COCA's authoring tools, and to discover if attitudes towards using AI in the classroom had been improved as a result of using COCA.

The university teachers were familiar with ITSs although none of the subjects was experienced in using COCA. Some of the school teachers had used computers before and some had no computer experience.

2 Background

KAFITS (Murray & Woolf, 1992), is a set of authoring tools enabling teachers to build ITSs directly. The tools allow domain knowledge and tutoring strategy knowledge to be created and edited. KAFITS facilitates teacher involvement in the ITS design process, alongside the knowledge engineer. COCA, whose aim has been to allow the teacher to work alone with the system, is necessarily a less complex system. KAFITS has been evaluated in a case study with one high school teacher and two education graduates (Murray, 1993). This was very much a qualitative evaluation but nonetheless gave some encouraging results in terms of productivity, with about 85 hours of effort needed for an hour of instruction.

Although there have not been any other evaluations of ITS construction tools, a certain amount of work has been done on suitable evaluation for ITSs, which naturally concentrates more on the use of the system by students rather than teachers. Different evaluation techniques have been studied (Mark & Greer, 1993) in order to point towards evaluation methodologies for ITSs. This work is relevant to COCA as it discusses the evaluation of different aspects of an ITS architecture, and most notably, the teaching strategy. It suggests changing the teaching strategy while keeping other components fixed as a possible means of evaluating effectiveness. Of all the techniques discussed, formal experimental techniques are put forward as suitable for summative evaluation of different components of the system.

Educational impact (Littman & Soloway, 1988) is a major test of an ITS in terms of both achievement and effect, expressed by how well a student learns and what attitudes are retained after that learning. Although we are not evaluating the ITSs built with COCA, we can still use these criteria with teachers to see whether they can use COCA's tools successfully, and also what effect this has on their attitude towards using AI tools in the classroom. These ideas are reflected in the aims of this evaluation of COCA.

2.1 General evaluation methods

A general discussion of different methods for evaluating software usability is given in Macleod (1992). One of the quickest methods of evaluation would be to give the system to teachers, asking them to use it as much as possible, and to write a report. This helps find important problems with the system, but is very dependent on the teacher. It is also similar to the case study reported in Major (1993a) and so would be of less value to us here.

Analytic evaluation methods are used early in the design process. They may employ a model of the user and apply it to a specification of the system. This model basically predicts the different interactions likely to be performed by the user. The advantage is very cost effective evaluation, done before implementation.

Controlled experimental studies of software usability are difficult because there are so many dependent variables when it comes to using a computer, such as the users themselves and their attitudes, the tasks they perform and the different environments in which they work. The standard procedure is to separate groups of subjects and to get them to perform the same experiment with one crucial variable changed. The results of the different groups are used to prove or disprove hypotheses about that variable. But isolating one such feature in COCA's context would be a very difficult exercise. An assessment of the value of controlled experiments with regard to evaluating software is given in Monk (1985).

Survey methods (Chin, Diehl & Norman, 1988) are the cheapest form of obtaining evaluation information. They are usually done by questionnaire. Quantitative data can be obtained if more than ten subjects are used, as well as qualitative data. Such a method would certainly be applicable with COCA, as long as the questions are directed towards the aims of the study.

A final set of methods are known as observational methods. This involves watching users working with the system, either directly or by the use of video, and recording the process. This information can then be analysed. This is more costly than using a questionnaire, but allows better access to the task itself. One such technique incorporates a think-aloud protocol (Wright, Monk & Carey, 1991).

The methods most applicable to COCA are survey and observational methods. The survey can be used by asking direct questions about COCA's usability and sufficiency, and about attitudes towards rules for representing teaching strategies. COCA allows users to save their knowledge bases and these can be analysed to allow observational evaluation. Further users could record their interaction with COCA by noting all the decisions they make and the functions they use when performing a task.

3 Experimental aims

The primary aim of the experiment was to assess whether or not COCA will be useful to teachers and authors. A number of issues were considered which are numbered below:

1 A first issue was the teacher's attitude towards AI techniques as opposed to traditional computer-assisted learning. The experiment investigated the teachers' perceptions of the changes in their attitude. It should be noted that this is a subjective measure. Although it could be suggested that the teachers' perception of their attitude change will always increase with an environment change, we observed in Major (1993a) that a system not offering control over teaching decisions would have been a serious disincentive to teachers with respect to using AI in the classroom.

2 Another question of particular interest was whether teachers considered they had had a conscious view of their strategies before using COCA. This is also a subjective measure.

3 Further, the experiment set out to discover whether teachers felt that the heuristics provided by COCA were likely to be sufficient to control teaching behaviour.

4 At a more practical level, the experiment set out to see which of COCA's facilities in the opinion of the teachers, were easy to use, and which were little used and difficult. Further, it attempted to assess which areas of authoring support subjects felt were likely to have to be extended in future versions. This is important when considering how future research can best build on COCA.

4 Evaluation method

The material chosen for the task was the domain of the American Revolution, which together with its associated strategy and meta-strategy, formed part of a conference demonstration of COCA (Major and Reichgelt, 1992). This teaching material, which is largely declarative in nature, makes use of a broad range of the features of COCA, including user-defined attributes, non-computer based tasks, hints and meta-strategic rules, and was constructed with the help of a history teacher who did not carry out the experiment. The fact that the subjects did not have any historical expertise is not relevant as we are testing COCA-1's authoring tools and not the final ITS produced.

There were 10 subjects consisting of 5 university teachers and 5 school teachers. Everyone completed the task and the questionnaire, but only 4 school teachers and 3 university teachers returned knowledge bases. The school teachers were all new to COCA and had not been used to find the requirements in Major (1993a). They were less experienced with computers than the university teachers. The subjects had received no training with COCA except for 3 school teachers, who had had 2 hours training together.

The subjects were presented with the domain material, initial teaching strategy and meta-strategy, in the form of prepared knowledge bases, together with a set of instructions in English for altering the teaching behaviour of the final system and extending the domain material. These changes involved creating rules, editing existing rules and using the meta-strategy. As they carried out the task, subjects were asked to write down each step that they took, including mistakes, using some examples on the instruction sheet to guide them as to what level of detail was required. The purpose of this interaction record was to give an indication of the complexity of the task. It would give some idea of how long the task took and how many mistakes were made. The task had been performed by the author, thus giving a measure of solution efficiency by comparing the number of rules to achieve given results.

Once the task had been performed, the subjects were asked to return their saved knowledge bases for analysis. They were also asked to fill in a questionnaire about the general use of COCA, the suitability of the teaching strategy representation, the domain representation, and a number of questions about how the subjects' opinions of AI in general and COCA in particular had been changed as a result of the task. As many questions as possible were given a closed set of possible responses, with 100mm scales being widely used. The aim of the analysis was to determine whether or not COCA was useful to teachers and authors, how attitudes towards AI were affected by using COCA, how conscious teachers were of their own use of strategies and whether or not COCA's heuristics were sufficient for teachers. All of these points were specifically addressed in the questionnaire.

5 Results

5.1 Knowledge bases produced

The knowledge bases produced were tested and analysed to see how much of the task had been completed successfully and to assess solution efficiency, by comparing with an optimal solution. Table 1 shows the relative success in the task.

The mean success rate for the whole task was 68%, which shows that the subjects were reasonably successful but nonetheless had some difficulty using COCA-1's authoring tools. The most difficult task was moving the summary to appear at the end of the teaching about a particular object rather than at the beginning. This involved deleting a rule from a group of rules that the subjects otherwise did not use. Consequently the success with this task was poor. Also difficult was the use of user-defined attributes. This coupled with feedback from the questionnaire suggests that the interface to COCA's domain authoring tools was not satisfactory and shows how to improve future versions of COCA.

The first four subjects were all school teachers. Subject 2 in particular seems to have needed considerable training with the system before being able to make any reasonable progress with it. Those subjects with the least computer experience did least well, which is an expected result. The time taken by each subject varied between 2 and 3 hours. Notes taken by the subjects while doing the experiment showed that there were few repeated attempts to perform a particular task, but rather once an initial attempt had been made for a task, it was usually abandoned.

Table 2 gives us some idea of the relative efficiency of the solutions produced by the subjects. This is achieved by comparing subjects' solutions with an optimal solution which solves the tasks making the minimum number of changes. It should be noted that the optimal number of changes is only optimal in terms of efficient editing, and makes no claim as to its efficiency as a teaching solution. The number of changes to each knowledge base was calculated by counting one change for each addition, edit and deletion. The score shown in the table is the number of correctly completed tasks, as shown in table 1.

KB	Task Description	Subjects 1	2	3	4	5	6	7	$\bar{x}$	σ_{n-1}
Domain	Change name	Y	Y	Y	Y	Y	Y	Y		
	Add objects	Y	Y	Y	Y	Y	Y	Y		
	Place objects	Y	Y	Y	Y	Y	Y	Y		
	User attributes	N	N	N	Y	Y	Y	N		
Subtotal (%)		75	75	75	100	100	100	75	86	13
Strategy	Change name	Y	Y	Y	Y	Y	Y	Y		
	Stay after fail	Y	N	Y	Y	Y	Y	N		
	Fill in blank test	Y	N	Y	Y	Y	Y	Y		
	Categorise test	Y	Y	Y	Y	Y	Y	Y		
	Test on dates	Y	N	Y	Y	Y	N	Y		
	Move summary	N	N	N	N	N	N	N		
	Swap date/causes	Y	N	Y	N	Y	N	N		
	Teach user attributes	N	N	N	N	N	N	N		
Subtotal (%)		75	25	75	63	75	50	50	59	19
Meta-strategy	Change name	Y	Y	Y	Y	Y	Y	Y		
	Remove implications	Y	N	N	Y	Y	Y	N		
	Restore implications	Y	N	N	Y	N	Y	N		
Subtotal (%)		100	33	33	100	67	100	33	67	34
Total (%)		80	40	67	80	80	73	53	68	16
Score (max 15)		12	6	10	12	12	11	8	10.1	2.3

Table 1: Percentages of tasks achieved by each subject

The efficiency for each subject is then calculated as the number of completed tasks per change to the knowledge base made, relative to this number for an optimal solution.

	Subjects Optimal	1	2	3	4	5	6	7
Domain changes	4	3	3	3	4	4	4	3
Strategy changes	13	7	5	10	9	11	8	6
Meta-strategy changes	3	3	1	1	3	3	3	1
Total changes	20	13	9	14	16	18	15	10
Score	15	12	6	10	12	12	11	8
Efficiency $\frac{(\frac{Score}{Changes})}{(\frac{Score_{opt}}{Changes_{opt}})} \times 100\%$	100	123	89	95	100	89	98	107

Table 2: Relative efficiency for each subject

We can see that although we have only an approximation to the efficiency of the solutions, that they are all very close to the optimal solution. The reason that it is possible to be more efficient than the optimal solution is that the more difficult tasks, which are completed in the optimal solution, require more changes to the knowledge bases than the majority of the tasks performed by the subjects. So a subject who has not attempted the more difficult tasks will have made fewer changes in proportion to his score than the optimal solution, and thus will have a higher efficiency. Subject 2 scored relatively highly despite not achieving many of the tasks. This was because the tasks she did achieve were done efficiently, and she did not make many attempts at those she could not manage. A lower efficiency also shows that a subject has taken longer to perform the task in comparison to what was achieved. Generally, we can see that the subjects managed to make the changes to the knowledge bases without using more rules than necessary, and so that COCA-1's flexibility is not at a high cost in terms of long editing sessions.

5.2 Questionnaire findings

Some of the questions asked for a mark on a 100mm scale, whereas others asked for more subjective opinions about teaching strategies and the use of AI in general. Table 3 gives the results. The scores are given with 0 as the worst possible score and 100 as the best. The average and standard deviation are also given. The subject numbers correspond to the numbers in tables 1 and 2 and are included for reference.

The first 5 subjects are school teachers and the rest university teachers. Firstly, and most generally, we can see that the majority of averages are over 50, putting them in the positive half of the scale as far as COCA is concerned. Question 3.3, regarding the use of variables in rules, shows the largest standard deviation. This results from the university teachers suggesting that they were strongly in favour of variables, whereas school teachers felt they would be highly unlikely to use them. On the other hand, question 3.8, regarding the practicality of building libraries of teaching strategies with COCA for general use, had the lowest standard deviation and the highest degree of consensus. This shows the potential of a system like COCA within a school or other teaching environment. Other points to notice are the high

Qu.	Description	Subjects										$\bar{x}$	σ_{n-1}
			1	2	3	4	5	6		7			
1.2	Terminology clear	35	50	27	44	74	88	85	94	75	72	64	24
1.3	Confident usage	41	40	6	40	31	61	86	22	82	55	46	25
1.4	Getting lost	57	17	75	44	64	68	75	61	52	46	56	17
1.5	Broad usage	–	87	50	30	51	89	85	95	23	73	65	27
1.6	UI easy	34	51	36	37	70	21	83	93	66	90	58	26
1.7	Tools useful	79	51	31	54	95	90	83	94	97	86	76	23
2.3	Domain tree clear	80	38	56	43	96	100	99	96	15	94	72	31
2.7	Course libraries	–	–	92	69	89	–	78	32	93	96	78	23
2.8	Domain distortion	–	–	–	29	87	–	–	67	96	17	59	35
3.3	Likely to use variables	64	51	3	58	4	98	100	93	–	83	62	37
3.6	Suits your teaching	70	81	95	66	89	30	48	44	93	–	68	23
3.7	Decisions easily made	24	80	–	63	92	11	87	26	92	86	62	33
3.8	Strategy libraries	–	89	99	81	93	93	98	86	94	83	91	6

Table 3: Answers to questions using a 100mm scale

scores, particularly amongst the school teachers, to question 3.6, regarding the completeness of COCA's teaching strategy model with respect to their own teaching style. The only score to be strongly negative was that of question 1.3, regarding whether users often felt lost in the system. It is clear that COCA's user interface needs to be improved to increase users' confidence and stop them getting lost.

The next part of the questionnaire measured changes of attitude towards teaching strategies and AI in the classroom. We shall concentrate on our school teachers as it is their attitudes that are perhaps more important. The two questions of interest to us are question 1.8, regarding the use of AI as opposed to CBT in the classroom, and question 3.10, which asked about attitudes towards teaching strategies. Also relevant are some of the comments made in the unstructured comments section of the questionnaire. Without exception, all our school teachers felt that COCA had shown them that AI could be useful in the classroom and had something to offer. They of course mentioned the problems of the interface and terminology, but could see the underlying usefulness of the system. With regard to teaching strategies, none of the school teachers had had as structured a view of strategies as COCA. Four of the five felt that forcing the user to consider their teaching in such a way was useful. None suggested that this view of teaching made authoring either difficult or distorted.

A number of other points were made in the responses to the other more open questions. The main concerns with the system were its terminology, documentation, user interface, student model and domain representation. The first three of these would be the main effort in any development of COCA's authoring tools for commercial use. Although the student model is weak it could easily be extended to allow any teacher-defined attribute (psychological/pedagogical) to be given to a student and to control the teaching. The domain representation could also be extended, with perhaps the strategy interpreter giving the domain interpreter more control, thus allowing larger and more interesting pieces of domain material to be taught. Some other strengths of COCA that subjects mentioned were its potential for formalising teaching through the strategy model, and thus to become a useful tool for trainee teachers. Subjects thought that being forced to consider teaching decisions beforehand would be profitable, and would help very much with the mixed-ability problem in the classroom. Finally they felt that the ability to change the student rating thresholds (i.e. the student modelling behaviour) was very useful.

6 Concluding discussion

It is certainly the case that COCA-1 is a complex system, yet despite this complexity, the experiment has shown that the ideas and structure contained in COCA-1's authoring tools are worth carrying forward, although the tools themselves need to be made easier to use. The experiment has also provided insight into the attitudes of teachers towards using AI, and in particular the use of teaching strategies.

Although the number of subjects was not large enough to make any categorical statements, it does give us a basis upon which to examine the different attitudes towards COCA, and also to examine differences between subjects. Those subjects who were university teachers were typically much more experienced in the use of computers and were familiar with AI techniques. Despite the breadth of tasks the majority of subjects completed well over half the tasks, with school teachers not doing particularly worse than university teachers. The high standard deviation on the meta-strategy tasks does suggest that such a level of abstraction above the domain was difficult for some subjects. Further the efficiency of use was very good by all subjects. This is an encouraging point especially for naive users.

In relation to the first experimental aim, COCA improved the perception of their attitude towards AI amongst the school teachers from neutral to positive. We can conclude then that the everyday use of tools like COCA for allowing AI models of teaching to be manipulated by teachers is likely to be a fruitful means of improving the acceptance of AI in schools. The university teachers opinion of their change in attitude was that it was not changed by the experiment in that they remained in favour of using AI

The second aim of the experiment concerned the subjects' opinions of their views on teaching strategies. Those subjects who did not feel that they had had an explicit concept of a strategy were typically the school teachers. They suggested that the formalising of a process that had previously been implicit was useful. The university teachers felt they had had a conscious view of strategies before.

With regard to the third aim, that of the sufficiency of COCA's strategy heuristics, the subjects said that COCA was indeed sufficient for their requirements. Further, there was strong support that building up libraries of strategies would be a practical way of using COCA. Those subjects who already had an idea of using strategies, typically university teachers, felt that COCA's model of a strategy and facilities for building strategies were sufficiently flexible for their needs. Indeed no subjects suggested aspects of the teaching process that were not catered for in COCA's tools.

The final aim was to discover those aspects of the system that were easy or difficult to use, and thus which aspects might need to be extended. With regard to this aim the experiment has also shown a number of weaknesses with COCA. We have already mentioned the domain representation and student modelling. Other points included the need for better and more comprehensive documentation and improvements in the interface. None of the subjects felt the teaching strategy model was a weakness. However, the fact that a number of tasks were not completed suggests that COCA's authoring tools need to be made more intuitive.

In summary we can say that COCA is usable enough for simple tutoring systems to be built, suggesting that a system like COCA should be pursued further, particularly with regard to the points arising from the final aim of the experiment. The school teachers who were suggesting the use of strategy and domain knowledge base libraries were highlighting a real problem for a system of COCA's type, namely that power is required at the strategy level, and yet the authoring task must be very simple if teachers are going to use a system on an everyday basis. To address this, a new version of COCA is under development. which very much simplifies the authoring task by hiding all rules from the teacher and giving graphical controls for strategy construction. This new version, running under Windows 3.1 on a PC, will be used as the basis of further experiments to investigate the nature of meta-strategic knowledge.

References

Bennett, N. (1976). *Teaching Styles and Pupil Progress.* London: Open Books.

Brna, P., Ohlsson, S. & Pain, H. (1993). Preface. In *AIED-93*, University of Edinburgh. AACE.

Chin, J.P., Diehl, V.A. & Norman, K.L. (1988). Development of an Instrument Measuring User Satisfaction of the Human-computer Interface. In *CHI-88*, 213-218. ACM.

Eaton, S. & Olson, J. (1986). "Doing computers?" - The micro in the elementary curriculum. *Journal of Curriculum Studies*, 18, 343–344.

Littman, D. & Soloway, E. (1988). Evaluating ITSs: The cognitive science perspective. In Polson, M. & Richardson J. (Eds.) *Foundations of intelligent tutoring systems.* Hillsdale, NJ: LEA.

Macleod, M. (1992). *Usability Now!* An Introduction to Usability Evaluation. National Physical Laboratory.

Major. N.P. & Reichgelt, H. (1991). Using COCA to build an intelligent tutoring system for simple algebra. *Intelligent Tutoring Media*, 2(3/4), 159–170.

Major, N.P. & Reichgelt, H. (1992). COCA - a shell for intelligent tutoring systems. In Frasson, C., Gauthier, G. & McCalla, G.I. (Eds.) *ITS 92, Lecture Notes in Computer Science, 608.* Berlin: Springer-Verlag.

Major, N.P. (1993). Teachers and Intelligent Tutoring Systems. In *The Seventh International PEG Conference*, Heriot-Watt University, Edinburgh.

Major, N.P. (1993). Reconstructing Teaching Strategies with COCA In *AIED-93*, University of Edinburgh. AACE.

Mark, M.A. & Greer, J.E. (1993). Evaluation Methodologies for Intelligent Tutoring Systems. *Journal of Artificial Intelligence and Education*, 4(2/3), 129–153.

Monk, A. (1985). How and when to collect behavioural data. In Monk, A. *Fundamentals of Human-Computer Interaction.* London: Academic Press.

Murray, T. & Woolf, B. (1992). Tools for Teacher Participation in ITS Design. In Frasson, C., Gauthier, G. & McCalla, G.I. (Eds.) *ITS 92, Lecture Notes in Computer Science, 608.* Berlin: Springer-Verlag.

Murray, T. (1993). Formative Qualitative Evaluation for "Exploratory" ITS Research. *Journal of Artificial Intelligence and Education*, 4(2/3), 179–207.

Plowden Report. (1967). Children and their Primary Schools. *Report of the Central Advisory Council for Education (England).* London: H.M.S.O.

Spensley, F., Elsom-Cook, M., Byerley, P., Brooks, P., Federici, M. & Scaroni, C (1990). Using multiple teaching strategies in an ITS. In C. Frasson and G. Gauthier (Eds.) *Intelligent Tutoring Systems: At the crossroads of Artificial Intelligence and Education.* Norwood, NJ: Ablex.

Wright, P.J., Monk, A.F. & Carey, T. (1991). Co-operative Evaluation: The York Manual. Department of Psychology, University of York.