



# The roles of models in Artificial Intelligence and Education research: a prospective view

Michael J. Baker

## ► To cite this version:

Michael J. Baker. The roles of models in Artificial Intelligence and Education research: a prospective view. Journal of Artificial Intelligence and Education, 2000, 11, pp.122-143. hal-00190395

**HAL Id: hal-00190395**

**<https://telearn.hal.science/hal-00190395>**

Submitted on 23 Nov 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## The roles of models in Artificial Intelligence and Education research: a prospective view

**Michael Baker** GRIC-COAST, CNRS & Université Lumière Lyon 2, 5 avenue Pierre Mendès-France, 69676 Bron Cedex, France, email: mbaker@univ-lyon2.fr

**Abstract.** In this paper I speculate on the near future of research in Artificial Intelligence and Education (AIED), on the basis of three uses of models of educational processes: models as scientific tools, models as components of educational artefacts, and models as bases for design of educational artefacts. In terms of the first role, I claim that the recent shift towards studying collaborative learning situations needs to be accompanied by an evolution of the types of theories and models that are used, beyond computational models of individual cognition. In terms of the second role, I propose that in order to integrate computer-based learning systems into schools, we need to 'open up' the curriculum to educational technology, 'open up' educational technologies to actors in educational systems and 'open up' those actors to the technology (i.e. by training them). In terms of the third role, I propose that models can be bases for design of educational technologies by providing design methodologies and system components, or by constraining the range of tools that are available for learners. In conclusion I propose that a defining characteristic of AIED research is that it is, or should be, concerned with all three roles of models, to a greater or lesser extent in each case.

*"Wenn wir an die Zukunft der Welt denken, so meinen wir immer den Ort, wo sie sein wird, wenn sie so weiter läuft, wie wir sie jetzt laufen sehen, und denken nicht, daß sie nicht gerade läuft, sondern in einer Kurve, und ihre Richtung sich konstant ändert."*

"When we think of the world's future, we always mean the destination it will reach if it keeps going in the direction we can see it going in now; it does not occur to us that its path is not a straight line but a curve, constantly changing direction."

Wittgenstein (1980), pp. 3 / 3e

### 1. INTRODUCTION

If, as some anthropologically-minded archaeologists would claim, the present is the key to the past, then perhaps the future is the key to the present? In this paper I assume the converse — that the present and the past are keys to the future — for the case of research in the field of Artificial Intelligence and/in Education (henceforth abbreviated to "AIED").

Any view of what objectives a research field may achieve in the future must be based on a view of the nature of the field in question, up to the present day. I characterise the past, the present and the near future of AIED research in terms of a combination of different roles played by models of educational processes, namely: models as scientific tools, models as components of educational artefacts, and models as bases for design of educational artefacts.

It should be noted that the views expressed here are not those of an objective historian of science, but rather of a researcher engaged in the field that is being discussed. In that case, description, prediction and prescription coincide to a certain extent.

One could say that there are basically three sorts of argumentative texts: those that argue (mostly) in favour of a particular view, those that argue (mostly) against one and those that

attempt to weigh pro and contra arguments in the balance (the conventional form of academic discourse)<sup>1</sup>. This text falls (mainly)<sup>2</sup> into the first category, and so no claim to exhaustivity is made in citing research that could constitute a rebuttal to the views argued for here. In the context of the special issue in which this text appears, I can only hope that some readers will be willing to supply counter-arguments and that a synthesis could emerge from any ensuing debate.

## 2. THE USES OF MODELS IN AIED RESEARCH

As with any field of scientific research, AIED involves elaborating *theories* and *models* with respect to a specific *experimental field*, in relation to the production of *artefacts*. What characterises a particular field is the nature of each of these elements and of the relations that are established between them: what types of theories are elaborated? what counts as a model? what is the experimental field studied? how close are the links between theories, models and artefacts? With respect to other research in the field of education, one of the specificities of AIED research lies in the different roles that models can play. A significant part of AIED research can be seen as *the use of computers to model aspects of educational situations that themselves involve the use of computers as educational artefacts, some of which may incorporate computational models*. By an educational situation I mean a situation that is designed in some way so that a specific form and content of learning will occur; by "educational process" I do not only mean the processes of learning and teaching, but also the larger scale processes by which social situations that are intended to enable teaching and learning to occur are designed.

There are thus three main roles for models of educational processes in AIED research, as shown in Figure 1.

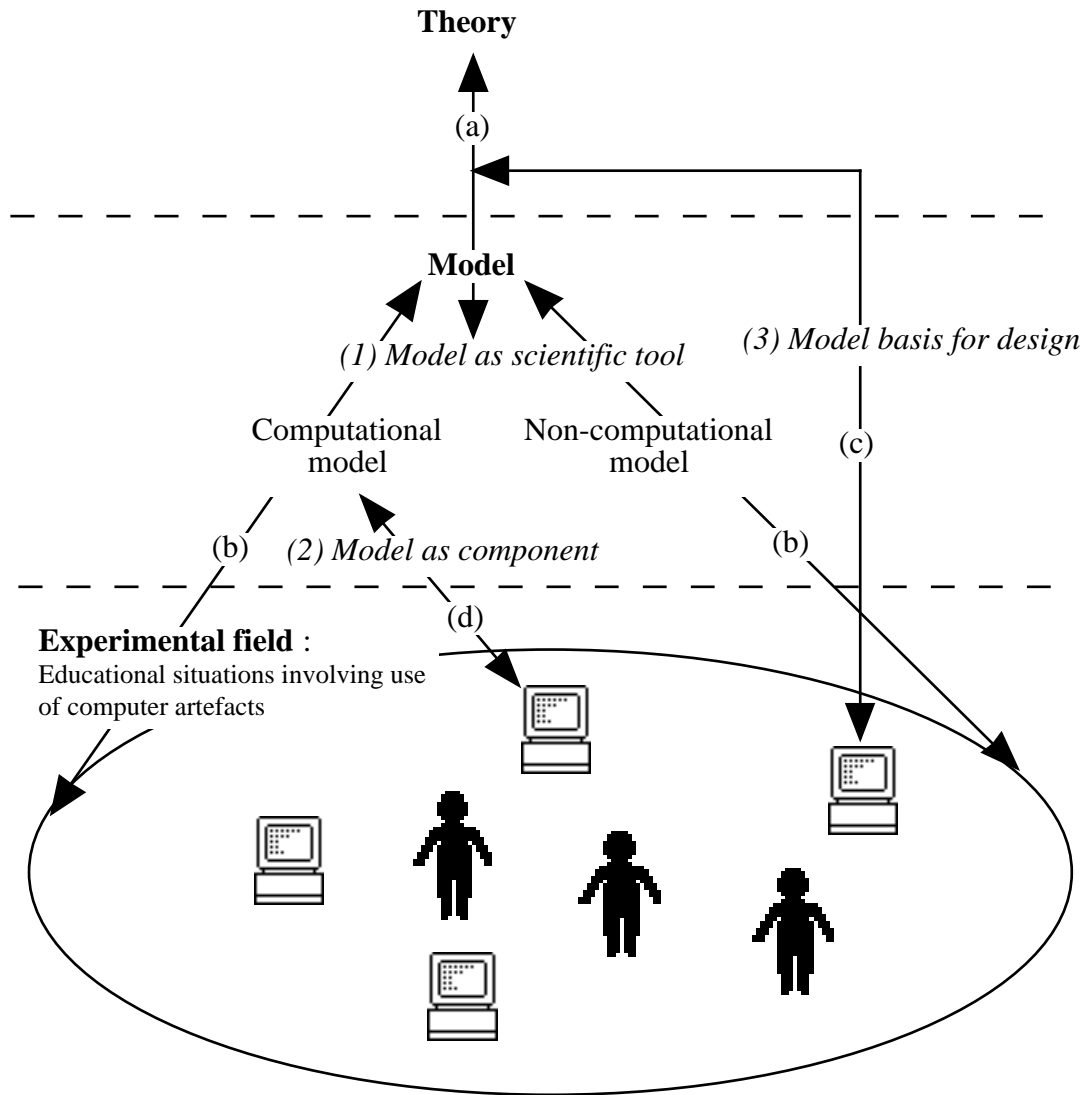
The three roles of models are as follows:

1. *Model as scientific tool*. A model — computational or other — is used as a means for understanding and predicting some aspect of an educational situation. For example, a computational model is developed in order to understand how the "self-explanation" effect works (VanLehn, Jones & Chi, 1992). This is often termed *cognitive modelling* (or *simulation*), although, as I discuss below, the term "cognitive" can have several interpretations.
2. *Model as component*. A computational model, corresponding to some aspect of the teaching or learning process, is used as a component of an educational artefact. For example, a computational/cognitive model of student problem solving is integrated into a computer-based learning environment as a *student model*. This enables the system to adapt its tutorial interventions to the learner's knowledge and skills. Alternatively, the model-component can be developed on the basis of existing AI techniques, and refined by empirical evaluation.
3. *Model as basis for design*. A model of an educational processes, with its attendant theory, forms the basis for design of a computer tool for education. For example, a model of task-oriented dialogue forms the basis of design and implementation of tools for computer-mediated communication between learners and teachers in a computer-supported collaborative learning environment (e.g. Baker & Lund, 1997). In this case, a computational model is not directly transposed into a system component.

---

<sup>1</sup> This is of course a simplification: there are also argumentative texts that aim to elaborate and define a point of view in the first place.

<sup>2</sup> The text does include some rejoinders to criticisms that have already been made by anonymous reviewers. These have been included mostly in the form of footnotes so as to preserve the general argumentative flow of the original text.



**Figure 1.** Three roles of models in AIED research<sup>†</sup>

<sup>†</sup> As is well known, one of the golden rules of such diagrams in (AI)ED research is that the meaning of the arrows must be clearly specified. The precise meaning of the arrows in the above diagram is in fact the principal subject of the rest of this paper. Briefly, and for the present, their meaning is as follows: (a) mutual adjustment between theory and model ; theory elaboration gives rise to model extension and vice-versa; (b) validation of a model with respect to an experimental field often leads to modification of the model ; (c) an educational system is 'based on' a theory/model (see § 5 below) ; evaluation of the system may lead to theory/model revision ; (d) incorporation of a computational model as a component of a computer-based educational system ; system evaluation leads to model extension.

Although researchers often attempt to establish a close relationship between 1 and 2 — e.g. cognitive-computational models of student problem-solving becoming student models in Intelligent Tutoring Systems (henceforth, "ITS(s)") — there is no necessary relation between the two, since it may be that the most effective functional component (in an engineering sense) of an educational artefact does not operate in a way that models human cognition.

These three possibilities are not, of course, mutually exclusive: most often, a given AIED research programme contains elements of each, to a greater or lesser degree. For example, one part of an educational system may be based on study of students' conceptions, and other parts may be based on using existing computer science techniques. However, it is not always possible to do this in a way that simultaneously satisfies requirements of each type of use of

models, i.e. produce a satisfactory scientific model that is an effective tutoring system component and which leads to an artefact that is genuinely useful in education. I believe that all three of these possibilities are valid and useful, provided that they are pursued in specific ways, that are coherent with the researcher's goals.

Before moving on to a discussion of the future of AIED research in terms of these three roles of models, I need to say something about what a model is<sup>3</sup>. Across different sciences, many different types of abstract constructions count as models — for example, descriptive, explanatory, analytic, qualitative, quantitative, symbolic, analogue, or other models. Without entering into an extended discussion in the philosophy of science, it is possible, and useful here, to identify a small number of quite general characteristics of models.

Firstly — and classically —, the function of a model is to *predict* the existence or future incidence of some set of phenomena, in a determinate experimental field. For example, models of stock exchange transactions should predict changes in financial indices; models of the weather should predict the weather tomorrow; a model of cooperative problem-solving should predict what forms of cooperation can exist (see below), and ideally what interactive learning mechanisms they trigger; a student model should predict the evolution of a student's knowledge states; and so on.

Secondly, a further, and just as important function of a model is to enable *elaboration* or *refinement* of the *theory* on which it is based, by rendering explicit its commitments on epistemological (what can be known and how?) and ontological (what is claimed to exist?) planes. It is generally accepted that there should be a link between the epistemology and the ontology: one should not posit the existence of entities without saying something about how they can be known. Such a relation between model and theory can lead to *explanation* of phenomena. A theory is not at all the same thing as a model; it consists of a set of quite general assumptions and laws — e.g. the views according to which human cognition is complex symbolic information processing, or that knowledge is a relation between societal subjects and the socially constituted material world — that are not themselves intended to be directly (in)validated (for that, the theory must engender a model). Theories are foundational elements of paradigms, along with shared problems and methods (Kuhn, 1962).

Thirdly, a model necessarily involves *abstraction from phenomena*, selection of objects and events, in its corresponding experimental field; it necessarily takes some phenomena into account but not others. It is not relevant to criticise a model as such by claiming that it does not take all phenomena into account, but one can criticise its degree of coverage of an experimental field. The modelling process itself involves complex matching processes during which objects and events are selected and structured so as to correspond to the model, within the constraints of its syntax. Tiberghien (1994) has termed this process one of establishing a *meaning*, or a *semantics*, for the model, in relation to its experimental field.

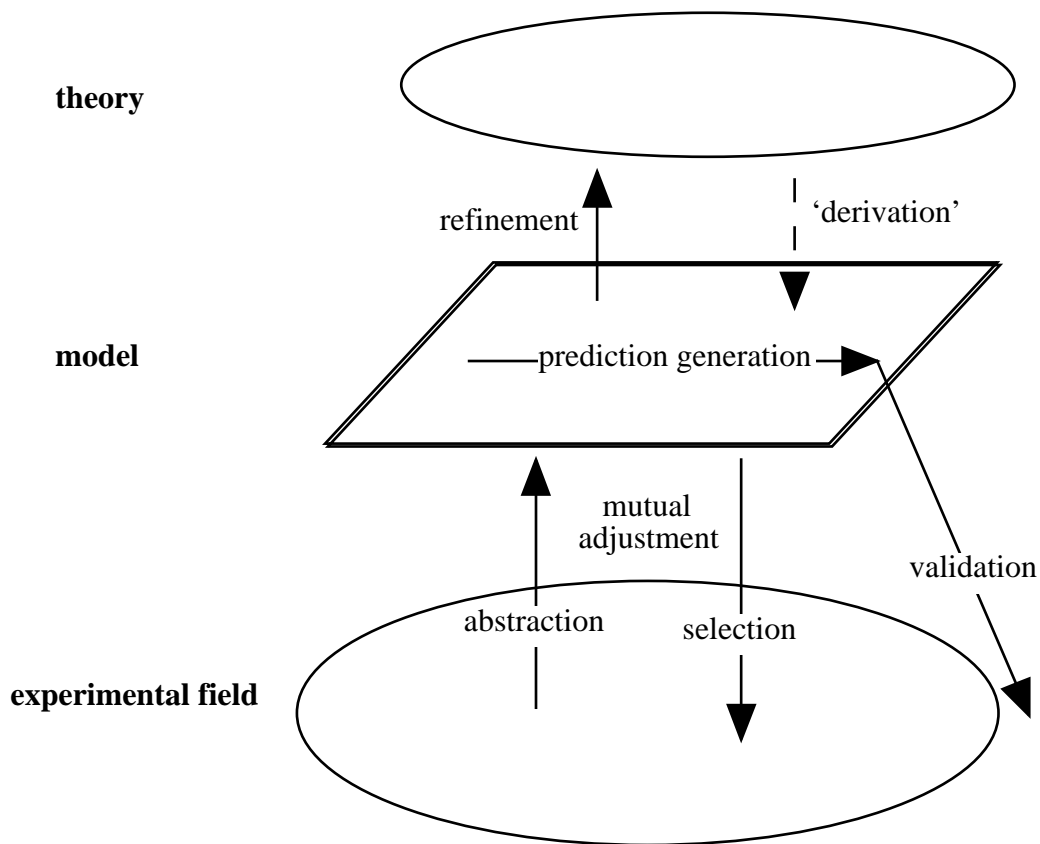
These general points are illustrated in Figure 2. I shall provide a simple example of a model of educational processes in the next section.

Where do artefacts enter into the picture? All research fields necessarily comprise aspects that are more or less close to the production and/or use of artefacts, in the sense of either 'applications' of theories or models, use of artefacts or instruments as experimental tools, sometimes on a large scale, or with respect to the study of artefacts themselves and their use, each of which can be a source of new research problems. Even highly theoretical work in mathematics, or descriptive work in botany, that is carried out as "pure research", may, perhaps decades later, find an unanticipated application via, for example, other domains such as physics or medical research. I do not believe that unidirectional 'application' exists: the relation between artefact, theory and model is always complex and multidirectional. Whilst it is clear that any field needs both theory and a close relation with the production of artefacts, it seems to me that one of the defining characteristics of AIED research is that it is closer to the theoretical end of the spectrum (cf. "computational mathematics" — Self, 1996). There is nothing intrinsically

---

<sup>3</sup> My views on the nature of theories and models in science have been very much influenced by the work of Tiberghien (1994, 1996) on learning to model in science. See also Greeno (1989).

wrong in that: for example, physics has for a long time comprised both theoretical and experimental branches. On that analogy, AIED research would be theoretically-oriented educational science, or even "Learning Science", that adopts a modelling approach.



**Figure 2.** General relations between theory, model and experimental field.

Despite this variety of *roles* and *types* of models, I think that AIED as a field nevertheless still largely operates with a somewhat restricted view of what models are — i.e. symbolic and computational information-processing models. Whilst this view has been important in defining the field as such up to the present, I do not think that it is fruitful or realistic as a unique 'model' for what the field currently is and will become. Other types of models of educational processes, that are not necessarily cognitive (in the above sense) nor computational in nature, can, and will I think/hope play an important role in AIED research, as I discuss below.

### 3. MODEL AS SCIENTIFIC TOOL

One aspect of AIED research involves elaborating formal or computational models of educational processes — learning, teaching dialogues, curriculum planning, etc. — as a means of understanding and predicting. I argue that AIED research would benefit from widening its theoretical foundations into cognitive science and from broadening its understanding of what a model is and is not.

#### 3.1 AIED models are not necessarily 'cognitive'

Although computational modelling of human thinking as a *symbolic information processing systems* has been much criticised over the past decade or so in cognitive science (for example,

from the point of view of various branches of situated cognition, situated learning and cognitive interactionism), I believe that modelling *per se*, of humans acting with computers in educational situations, will, in the medium or long term, become *more*, not less important. Let me make that claim more precise: criticism of the view of human beings as symbolic information processors is not the same as criticising (computational or other) modelling as such, but rather a specific type of computational model of human capacities.

It does seem, however, that the wholesale rejection of computational models of the symbolic kind, on the part of many researchers in cognitive science, can lead to a type of hyper-empiricism, that rejects modelling and formalisation as such. I am thinking of different varieties of ethnomethodological approaches to the study of learning situations that from the outset 'suspend' or 'postpone' theorisation, more or less indefinitely.

There is a different form of empiricism (or perhaps, of pragmatism?) in AIED research that does not consist in an attachment to 'the data', but rather in an emphasis on producing and evaluating "working systems". I have nothing to say against that view as such, if only to say that it is only *one* view of the role of models in AIED (the "model as system component" view, described below). As I already mentioned, all sciences need *some* more or less direct relation with production of artefacts; but they also need theories and models, otherwise it is difficult to embed artefacts in varied situations, and to understand why they are or are not effective in achieving their aims. From the point of view of the "model as scientific tool" approach, a computational model needs to be implemented to the extent that it can be validated. Once it no longer appears fruitful, in scientific terms, to extend that model further, it will no doubt gather dust once the theoretical and empirical insights it has yielded have been published.

I believe that modelling human activity in educational situations involving computers will continue to be an important scientific enterprise in the future, and especially in AIED research, for two types of reasons: socio-economic and scientific.

To begin with the socio-economic reasons, J.-F. Lyotard, in his book *The Postmodern Condition: A report on the status of knowledge*, states that:

"With the hegemony of informatics, a certain logic is imposed, and thus a set of prescriptions bearing on the utterances that are accepted as pertaining to 'knowledge'". (Lyotard, 1979, p. 13).

What he means is simply that computers are having such a growing influence in every aspect of life — including notably re-search for knowledge — that they will impose their standards on what counts as knowledge at all. Knowledge is, or will be, what is or can be formalised in a computational way. Let me be clear: I am stating no value judgement whatsoever on what I think should be the case, but simply stating what I think will be the case, in accordance with Lyotard's analysis. Naturally, such a powerful transformation of society will produce counter reactions of various forms, of which an emphasis on intuitive, intrinsically unformalisable or even situated knowledge is perhaps an example. I think that Lyotard's prediction, made around twenty years ago, has been largely validated, and so will continue to be true in the future. It seems rather clear that not only are computers influencing (invading?) almost every aspect of life — artistic, medical, economic, political, personal, interpersonal, ... — but they have also had a great influence on all areas of research, including the humanities. There is certainly no reason why this will be less so in the case of human sciences research in the field of education that is linked to computational modelling approaches, i.e. AIED research.

The second, scientific, reason is simply that the type of rigour and precision associated with modelling is what significant parts of most scientific endeavours seem to aim for. Certainly, the price to pay, as with any modelling approach (see above), is a degree of abstraction from reality, the leaving out of certain phenomena.

But if, as I predict, formal models of educational and other processes are here to stay, what type of models will they be, what will they be models *of*, what will their experimental fields be?

It is true that, since its beginnings, and up to the present day, AIED research has been largely wedded to the "Old Alliance"<sup>4</sup> between symbolic AI and information processing cognitive psychology. Amongst other things, this meant putting action and perception 'between parentheses'. Let me be clear at the outset: I think that there is *nothing inherently wrong* with such an approach; the number of insights into human thinking that it has produced is unquestionable. But the approach is fruitful *as far as it goes*, i.e. given the range of phenomena it can predict and explain. Whilst it has been fruitful in understanding, for example, common errors that children make in subtraction, or evolution in performance whilst using devices such as computer keyboards, it is perhaps not the first approach one would think of for understanding the coordination between hand, eye, brain, pencil and paper required when a child learns to write, or in understanding violence in schools, or in understanding how teachers' discourse in a specific classroom is determined by their social status as representatives of a particular profession (c.f. Wertsch, 1991), or ... . Whilst the information processing approach can of course yield some partial insights in these cases, perhaps few would deny that some significant phenomena would be left out; and it is legitimate for some researchers to want to include some of them in their research approaches.

There is no *a priori* reason why computational and symbolic models of socio-cognitive, of integrated cognition, perception and action, could not be elaborated. The question would simply be as to whether they were useful in solving important problems.

As is well known, there have been profound changes in cognitive science over the past two decades, changes that have put much greater emphasis on the interweaving of thinking, knowledge, social interaction and tool-use. If, as I believe, AIED research is firmly part of cognitive science research, then why have these changes in cognitive science had so little effect on AIED research?

I believe that it is and will be possible to have it both ways, steering a route between hyper-empiricism that eschews theorising and formalisation, and a form of idealism that produces theories and models with little anchoring in reality: i.e. retain a modelling approach whilst extending theoretical perspectives and the range of phenomena taken into account. If economists and (some) sociologists can develop formal and mathematical models of social processes, why should that not also be possible for AIED researchers, with respect to educational situations, broadly considered? Given that a purely cognitive model of teaching practice leaves out many significant phenomena, they why should it not be possible for models to be elaborated that take socio-institutional constraints into account, as well as the necessarily situated and reactive nature of the achievement of at least some of the teachers' goals in the classroom? I think that this is possible whilst at the same time being in complete agreement with Clancey's (1993) view that reality will always be more subtle than our attempts to understand it: that is true of any scientific research, and by definition, true of any attempt to model some part of reality.

An example of a theoretical approach that could be interesting for research in AIED<sup>5</sup> is *Cultural-Historical Activity Theory*, developed as a continuation of Leont'ev's work (1981), himself being a student of Vygotsky. The unit of analysis of this approach is not individual cognition, but rather "person(s)-acting-with-mediational-means, viewed as an indissociable whole" (Mammen, 1993). In other words, persons interacting with each other, with tools such

<sup>4</sup> The expression "Old [Aude] Alliance" refers to a longstanding relationship between Scotland and France that dates from a treaty of mutual military assistance against their common enemy — England — drafted by John Balliol of Scotland and Philip IV of France, and ratified on the 24 February 1296. I am grateful to Paul Brna for pointing this out to me.

<sup>5</sup> The work described in Baker, Hansen, Joiner & Traum (1998, 1999) constitutes a small step in this direction to the extent that it aims to propose a model for computer-mediated collaborative problem-solving that incorporates insights on the role of tools from Activity Theory with work on grounding carried out in AI and language sciences. Dillenbourg's earlier research (Dillenbourg, 1991; Dillenbourg & Self, 1992) on modelling socio-cognitive interaction can also be seen as a step in this direction to the extent that it is based on Vygotsky's notion that forms of discourse can be internalised as schemas for reasoning.



as computers, symbolic or natural languages, in the context of their culture. Is there any *a priori* reason why a — formal or other — modelling approach could not be based on such a unit of analysis? It really does make a difference to the whole structure of an activity such as collaborative writing whether it is achieved with a computer or with paper and pencil. But taking this into account, as most researchers do of course, should transform the theoretical perspectives that we use: the tools are not add-ons to a pre-formed activity, they constitute it. Certainly, the abstraction known as "mental representation" in some models would have a different ontological status, a different semantics, since according to this approach "consciousness" is a constitutive aspect of human activity, that has developed for reasons relating to the societal structure of activity — e.g. division of labour — over the history of cultures. Producing formal models of persons acting with auxiliary means would not necessarily produce results that would be acceptable to researchers in Activity Theory themselves, since the meaning of their theoretical principles would necessarily suffer transformation, in their transposition into the field of AIED. But I think that such creative transformations are legitimate, possible, desirable and probable.

Finally, it is true that over the past decade, there has been a shift in AIED research from studying individual learning and adaptive tutoring towards studying collaborative learning situations in distributed systems. But in that case, this shift of the object of study should be accompanied by a theoretical and methodological shift. In other words, for many research problems associated with the shift towards collaborative learning, the problems can not be appropriately solved whilst retaining theoretical approaches that assume a separation between cognitive and social processes (cf. Perret-Clermont, Perret & Bell, 1991; Dillenbourg, 1999). If researchers attempt to analyse collaboration as a conjunction of individual cognitions then, I would argue, they will run into serious methodological and theoretical difficulties. For example, trying to understand collaboration in dialogue as a conjunction of individual utterances or contributions faces the problem of the collaborative construction of meaning. Individual utterances, considered in isolation from their dialogue contexts, can not be assigned their 'full' meaning, or the meaning that is commonly understood and negotiated by the interlocutors themselves. This is most readily apparent in the limiting case of contributions produced by collaborative completion (Roschelle & Teasley, 1995). The primary unit of analysis is the exchange, not the individual move. Speech acts, and other aspects of meaning, are joint constructions, and should be analysed as such, using appropriate theoretical tools

In brief, my claim is that the theoretical foundations and methodological approaches of AIED research must evolve with changes in the types of situations that it aims to study.

### 3.2 AIED models are not necessarily computational

I believe that *computational models are not the only valid forms of models in AIED research or in any other form of research*. What is required in order for something to count as a model is minimally (see §2 above) that it possesses an internal coherence, that it exists in a specific (semantic) relation to a specific experimental field (the relation is complex, involving abstraction and prediction) and that it is possible to derive predictions from it that can be validated<sup>6</sup>. It is probably impossible for me to make this point without giving an example. The example I give is from modelling collaborative problem-solving in learning situations (Baker, 1994, 1995; Baker, in press-a).

The experimental field of this simple model is quite wide or general — dyadic collaborative problem-solving in situations designed to promote learning — and is represented initially by a finite set of interaction corpora, for different types of problems. Future research aims to extend this field in a systematic way.

Modelling begins with some simple observations (that are of course 'theory laden'). Three main forms of cooperation-in-interaction can be observed:

---

<sup>6</sup> Here I would not wish to restrict internal coherence and prediction generation to logical coherence and logical derivation since these are not the only forms of reasoning that can be used.

1. *co-construction* — each student contributes more or less equally to proposing and refining solutions to the problem at hand;
2. *one-sided collaboration* — one of the learners proposes solutions that are criticised or elicited by another;
3. *argumentation dialogue* — there is explicit disagreement about a proposal or proposals, the learners defend and attack the proposal(s).

A further observation is that the students may be 'working together' more or less well; they may understand each other's proposals to a greater or lesser extent (Edmondson, 1981; Clark & Schaefer, 1989), and their activity may be more or less 'in phase' (Baker & Bielaczyc, 1995).

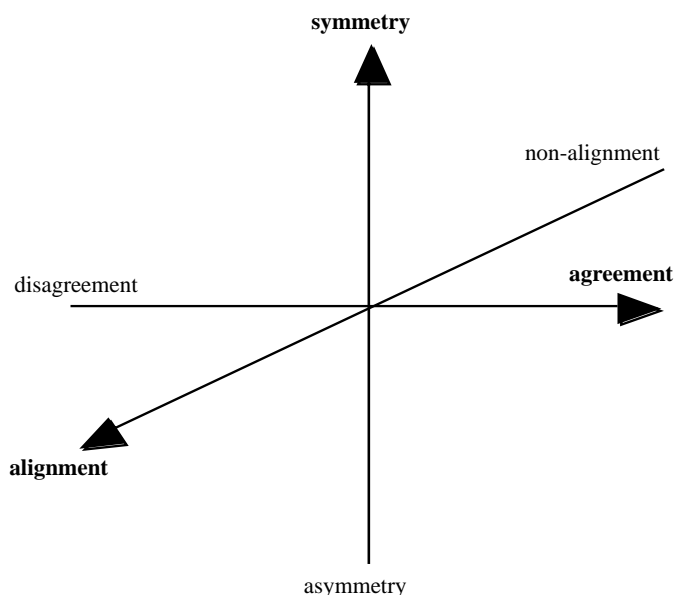
So far, modelling has consisted in a limited abstraction from the available data on the basis of a wide body of theory. The next step involves further abstraction and generalisation from these three abstracted categories, as follows.

Three fundamental dimensions underlie the above three categories, together with the additional observation:

1. degree of (*a*)*symmetry* of roles adopted in cooperative problem-solving,
2. degree of (dis)agreement,
3. degree of *alignment* (of mutual understandings, of students' problem solving stages).

These dimensions are abstractions: they do not directly 'correspond' to the data. The specific form "co-construction" is, for example, a type of interaction that is relatively symmetrical in terms of contributions to problem solutions, and in which the students are globally agreed. Figure 3 shows these three dimensions.

It is important to realise that the unit of cognitive analysis here is not the individual, but the dyad. For example, it makes no sense to say that an individual has adopted a symmetrical role, or is 'in agreement' (although it would of course be possible to take a model for group cognition as a basis for elaborating a dialogical theory of mind).

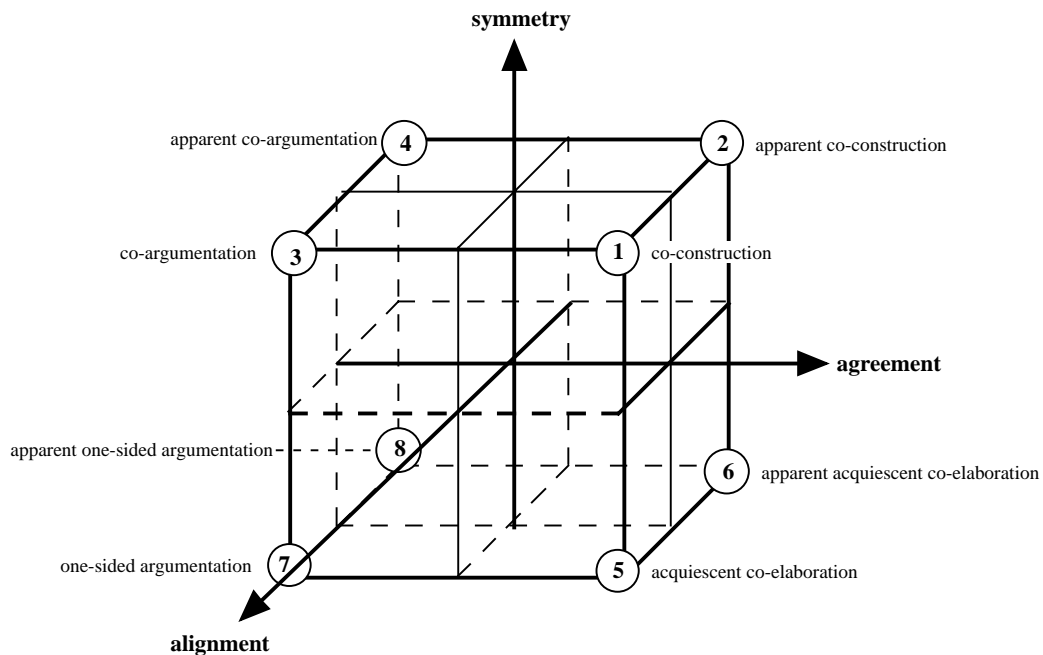


**Figure 3.** Three fundamental dimensions of cooperative problem-solving.

What makes this apparently simple diagram a 'model', in the strict sense of the term, is the fact that once these three dimensions have been abstracted<sup>7</sup> from three finite forms of

<sup>7</sup> The dimensions were of course abstracted under the influence of existing theories of dialogue, in this case notably those described by Bunt, 1989, Clark & Schaefer, 1989, and in my own previous work, Baker, 1994, 1995.

interaction, is that a very simple model syntax can be applied to *generate* and *predict* the existence of five other possible forms of interaction, to give eight *forms of cooperation* in total, as shown in Figure 4.



**Figure 4.** Eight basic forms of cooperation in cooperative problem-solving activity.

For example, there can be various forms of argumentation, that are more or less symmetrical (either each participant has an alternative proposal, or else one participant simply contests another's proposal), and more or less aligned (e.g. "arguing past each other"<sup>8</sup>). Once these forms are predicted by the model, then we can attempt to *look* for them<sup>8</sup>, in order to validate the model itself. The model's theoretical interest lies in the fact that it claims to reveal the operation of simple dimensions, 'behind' the complexity of reality. I think that that very simplicity is a virtue. The analytical model is elaborated further in terms of operational definitions of the three dimensions, and by specification of the *processes* that lead to the forms of cooperation. For example, symmetry is defined in terms of aggregates of roles, which in turn break down into specific actions (communicative or non-communicative), and the media (cf. Dillenbourg & Traum, 1996; Clark & Brennan, 1991) with which they are achieved. Disagreement is defined in terms of a theory and model of argumentation dialogue (Barth & Krabbe, 1982; Baker, 1996, Baker, in press-b) and a theory of linguistic feedback (Allwood, Nivre & Ahlsén, 1991). In relation to my earlier remarks on activity theory, a crucial aspect of what actual forms of cooperation occur concerns the material tools that are provided, especially in the case of computer-mediated communication.

I do not want to go into more detail on this model here. I simply want to make the following three points:

1. What is shown in Figures 3 and 4, with its attendant theoretical framework and operational definitions of the dimensions, *is a model*, albeit a relatively simple one, of a process that occurs in an educational situation. It would probably be best described as an *analytical model*, since it does not — at this level of detail — describe the cognitive-interactive *processes* that lead to the modelled phenomena (cf. Baker, 1994, 1995). It is a model since it exists in a particular semantic relation (abstraction, selection) to an experimental field, the model has a very simple syntax (generation of logical

<sup>8</sup> This does of course require an operational subtheory of argumentation that I shall not present in detail here.

possibilities of combinations of dimensions) and because it predicts the existence of forms of cooperation that were not initially observed.

2. *The model is not computational* and is yet, I would argue, useful in AIED research. For example — and as I describe below — it may be useful for design of collaborative educational technologies. It could also give rise to specific and local computational models (cf. Inaba & Okamoto, 1997) in human-computer educational dialogues, or, perhaps, be useful in interpreting experimental results, beyond (?) the field of AIED.
3. The model is *not 'cognitive'*, under a specific interpretation of that term that refers purely to individual cognition: the unit of analysis is the interaction sequence, the dyad's activity. I would describe it as being cognitive-interactional, and would be happy to abbreviate that to just "cognitive", bearing in mind my remarks on what that means here.

I sincerely wonder whether, for many types of problems that AIED research addresses, simpler models than those often proposed — e.g. those based on formally elegant epistemic logics — would not in fact be equally adequate to the scientific or design task in hand. The crucial issue, to my mind, concerns the question as to why one is attempting to model at all. And once researchers have answered that question in specific cases, it may not always be necessary to adhere to the current formalisms of the day as they exist in artificial intelligence research.

Let me try to summarise and simplify my view of the present and future of AIED research, in terms of the role of models as scientific tools. I think that modelling as such is necessarily here to stay in AIED research, as it is in any field of scientific research — perhaps no-one would disagree with that? However, the models involved are not, and will not be, necessarily restricted to symbolic information-processing models, and are not necessarily computational in nature, given that their purpose is to enable us to understand and predict educational processes, and perhaps to form good bases for design of educational artefacts. As a purely personal statement, I would say that AIED needs to — and I think that it inevitably will — open-out its theoretical horizons and anchor itself more firmly in cognitive science, considered in its broadest sense. There is no need to throw out the baby with the bathwater: the fact that certain computational models in AIED do not take some phenomena into account does not invalidate the point and value of modelling as such in AIED or any other field of research.

#### 4. MODEL AS COMPONENT OF AN EDUCATIONAL ARTEFACT

In terms of this particular current of AIED research, a computational model corresponding to some human cognitive ability is used as a component in an AIED system — Intelligent Tutoring System (ITS), Intelligent Learning Environment (ILE), etc. The most well known example is of course to be found in research on student modelling where, in many cases, a computational model of student reasoning or problem-solving in a restricted knowledge domain is used as a component of an ITS that attempts to model the evolution of an individual students' problem-solving processes throughout an interaction between the human student and the ITS of which the student model is a component.

As is equally well-known, there are several different views as to the nature of such computational models-as-system-components. The first view is not so different from the one described above, under the "model as scientific tool" rubric: the model-component should correspond in some way to human performance and, in the ideal case, should be able to predict the changes in the experimental field that would result from each tutorial intervention. For example, the model-component should be able to predict changes in the student's cognitive states that result from providing some specific knowledge. The technical problems associated with such a research approach are not our main concern here, although it is clear that their extreme difficulty is an important barrier to the continuation of this line of research. A second

view is that a model-component does not necessarily have to predict human performance, nor correspond to it in any direct or indirect way (for example, it does not have to operate within the constraints of human cognitive limitations). Thus the model-component is precisely a functional component of a tutoring system architecture, whose value is to be *evaluated* (i.e. not "validated") in terms of more global criteria of success — to what extent does addition of the model-component to the architecture lead to improved learning?

Whichever of these two views concerning the status of model-components one adopts, it is clear that these components are intended to be parts of educational artefacts, i.e. objects constructed to fulfil some practical educational purposes in specific educational situations. So we have to ask what those artefacts might be, what the educational purposes might be, and in what educational situations those purposes can reasonably be pursued.

Let us begin with the most well-known case: computational models as components of ITS that are intended for the purpose of providing individualised instruction, in school learning situations. In this specific case, my view is as follows: *ITS, as independent providers of individualised instruction, will not be appropriated in the near future into school classrooms* (I mean primary to secondary level classrooms, involving learners from 5 to around 18 years old). At least three main questions arise from that statement, that I shall address briefly in turn:

- Why might not ITSs, as defined above, be appropriated into school classroom situations in the near future?
- What other types of educational situations might individualising ITSs be appropriate for?
- What other types of computational educational artefacts might be integrated into classrooms?

#### 4.1 ITS for individualised instruction and classrooms in the near future

There exist significant differences in the way in which education is organised in different countries. I shall therefore restrict my remarks to European countries, and France in particular. When I use the word "classroom" in the rest of this section, I mean primarily "French classroom" and secondarily "European classroom", usually for any age from 5 to 18 years, within obligatory education that is provided by the state. It may be that some of these remarks will also be relevant for other types of schooling and other countries.

Secondly, in this section, by the term "ITS" I mean a system that is designed for providing adapted instruction on an autonomous basis to an individual student. That will usually mean the use of a student model, pedagogical knowledge and appropriately represented domain knowledge in order to sustain an adaptive tutorial interaction, perhaps within the context of use of a learning environment<sup>9</sup>.

---

<sup>9</sup> An anonymous reviewer of this paper has already objected that if one (i.e. the present author) adopts such a narrow or "traditional" view of what an ITS is, then the views and arguments that follow are tantamount to knocking down a straw man (i.e. having deliberately established a weak version of a claim so that it can be more easily attacked). My preliminary reply to this objection is that, firstly, this characterisation of ITS is not a straw man since at the present time, these types of ITS are the ones that are most developed technologically, and thus the most likely candidates for putting into the classroom ; the number of researchers who, from their writings, imply that their ITS as described above, *could* be used in the classroom is larger than zero. Secondly — and as I discuss below — I do recognise that some very recent developments in AIED systems (such as systems that are open to teachers, that allow inspection of their internal states, and so on) do go in the direction of addressing some of the problems I raise here. But I think it unlikely that these systems will be robust enough within the next ten years so as to have wide use in schools. At the time of writing, in the schools in which I have worked, the use of educational technology on the few computers that are available (typically, three or four computers in the school library, linked to internet via a single telephone line, plus a few more computers in science laboratories) is restricted to the use of educational CD-ROMs, some numerical simulations in science class and searching for information on the web.

I do not believe that ITS for individualised instruction will be appropriated into schools in the near future because of the problems of *social responsibility and legitimacy*. It is a truism that schools are organised so that children will learn; and what they learn (e.g. mathematics, technical drawing, music, citizenship skills, how to work with others, etc.) and how they do it is potentially the concern of every citizen, every socially organised structure (such as private companies, hospitals, theatres, ...) and thus of democratically elected governments. Certain persons — school administrators, directors, teachers — are employed by the state or other collective bodies in order to ensure that the *socially legitimised* form and content of learning takes place, and they are thus *responsible* for the effective carrying out of those objectives, however they are defined, with whatever degree of precision. If there is to be some degree of *devolution* of that social responsibility towards computer systems, then, since computers themselves are not socially responsible actors, in the final analysis the teachers, the school directors, the government ministers, will have to assume that responsibility. It makes no fundamental difference if we propose, in a futuristic manner, that the teacher's function can and will be modified from "content provider" to "curriculum manager": they are still ultimately responsible.

Now, if a teacher, for example, is to accept devolution of part of responsibility for teaching to a machine, that individualises its instruction, then not only will the teacher have to manage the individualisation within a group (such as a class), but the teacher will also have to *understand* how that individualisation occurs in order to *accept* the devolution of responsibility. Software producers' manuals and demonstrations are unlikely to be sufficient in this respect; no doubt the system will have to be 'transparent', in some sense of the term, for teachers. This is one of the classic problems that faced expert systems. Imagine a room full of students each working at their individualising machines, or even a group of learners working at a distance, with internet contact with teachers. At some point a learner is likely to ask "every time I do/say x it does/says y (to me); why does it do that? I don't understand what it's getting at, what it wants from me?". This could occur however sophisticated the system's explanations are: at some point explanations come to an end. Now the teacher must be able to provide some kind of answer to the student, and that requires specific knowledge, and usually that requires *teacher training and experience*.

If one thinks that devolution of responsibility for teaching from teachers towards *computers* is a straightforward matter, then one only has to remember the common problems that can be associated with devolution of responsibility to *parents*, who often teach their children certain skills (e.g. reading, basic arithmetic, tennis) out of school time. Every parent has to at least consider the question of how what they teach to their child at home will relate to how the child will learn that subject in school with the teacher; some form of more or less direct cooperation is required between parent and teacher.

So my conclusion is simple: individualised instruction by an ITS, assuming it is feasible and desirable, will not be appropriated in schools unless its decisions can be readily comprehensible by teachers and other social actors in the educational system<sup>10</sup>. ITS systems need to consider the needs of not only students, but also of teachers. There is perhaps a paradox, which is that in practical terms, over the next ten years at least, it is not certain that the

---

<sup>10</sup> I am grateful to an anonymous reviewer for having pointed out to me that much the same problems of understanding, teacher acceptance and devolution of responsibility could also arise with "unintelligent" CAL or Integrated Learning Environments. For example, neither student nor teacher might understand why some CD-ROM courseware obliges the student to move from one screen to the next, or how a video clip is supposed to relate to a textual explanation. However, in the two cases — courseware / ITS — there is surely a difference here in the degree of complexity of understanding required. In the case of the courseware, student and teacher may be led to reconstruct the courseware designer's intentions in juxtaposing information screens, whereas in the case of an ITS system they may be led to do that in addition to attempting to understand the way that the machine 'reasons'. And in the latter case, the reasoning may not be open to inspection in every case (*pace* systems that do attempt to explain their reasoning and "inspectable" or "open" student models — e.g. Bull & Broady, 1997 ; Dimitrova, Self & Brna, 1999).

kinds of systems that will be open to teachers as well as students, and comprehensible to both, will be excessively complex, from the point of view of the leading edge of AIED research. Yet that is, I think, normal: in any research field there is a more or less wide gap between research and practical application. However, just as the petrochemical industry ultimately and very indirectly relies on discoveries in theoretical and experimental chemistry, and the latter is stimulated by new problems and techniques developed in the former, I think that the same is true for education. There is no reason at all to be ashamed of carrying out theoretical research in AIED, provided that that is not the only form of research that exists.

Putting ITSs into schools is not just a question of researchers designing systems, with or without the collaboration of specific actors (students, teachers etc.) in the educational system. It is also, and perhaps above all, a matter that has to be addressed on *structural, institutional, socio-economic* and *political* levels — adapting curricula, transforming the nature of teaching and teacher training. And that takes time. I do understand researchers (like myself) who say that such things are not their direct concerns, as researchers *per se*. But in that case, if researchers are not prepared to engage to some extent in these social and institutional issues, then I do not think it is advisable for them to seriously envisage putting ITS or any other systems into schools.

In fact, to be even more brutally realistic, the question of ITS systems, or any other systems for that matter, being in schools or not is not yet really on the agenda: it is certainly not the case that all schools and classrooms even possess computers, and it is certainly not the case that all teachers have been trained in managing teaching situations involving computers. We can realistically expect that this will be the case within the next ten years, which makes this a crucial period for establishing the foundations for integration of educational technologies in schools, but perhaps not for their integration as such.

I should make it clear that I am *not* claiming that putting ITSs into schools is impossible or even ill-founded, but simply trying to identify some of the major problems associated with such an enterprise, over at least the next ten years. The research described by Koedinger and colleagues (Koedinger, Anderson, Hadley & Mark, 1997) is an excellent — and relatively rare — example of work on integrating ITS into schools that illustrates perfectly the degree of success that ITS can have today, ways in which teachers can be involved in integrating ITS, problems associated with that integration and the level of resources required. These authors describe a large-scale evaluation of the PAT ("PUMP Algebra Tutor", or "Practical Algebra Tutor") intelligent tutoring system for algebra, in high schools in the Pittsburgh (USA) area, that was carried out within the framework of the "PUMP" project ("Pittsburgh Urban Mathematics Project").

Firstly, it is clear that PAT was successful in teaching : 470 students in experimental classes, using PAT, outperformed students in comparison classes by 15% on standardised tests, and by 100% on tests targeting PUMP objectives.

Secondly, teachers were involved in the overall project in various ways : in developing a more accessible algebra curriculum, in designing new problems using an authoring environment, in choosing appropriate help messages as well as desirable next activities and, of course, in supervising the use of PAT in their classrooms. The authors report that teachers were generally enthusiastic about using PAT since it was able to deal with a large proportion of students' questions, which thus gave them more free time to give individualised help to students in difficulty.

With respect to my remarks above, three points can be made concerning the PUMP/PAT research. Firstly, some problems with overall classroom integration remained : "... participating teachers were not certain how to coordinate students' differing rates of progress through PAT lessons with the material being addressed in the regular classroom." (ibid., p. 37). Although, the authors do not discuss this point, I would speculate that in part, this problem may relate to difficulties that teachers may have on one hand in matching their own ideas about students' skills in algebra with the fine-grained way in which they are represented as production rules in PAT and, on the other hand, in understanding the Bayesian procedure for estimating students' strengths and weaknesses relative to those production rules. It is of course possible that many

teachers would be willing to 'live with' a degree of lack of understanding on their part of how an ITS works, provided it gives the right learning results, as PAT does. That is an empirical question that has not yet been answered, to my knowledge.

Secondly, whilst in the Pittsburgh area, teachers could produce " ... a more accessible algebra curriculum ... " (op. cit., p. 31) within the constraints of recommendations of a national organisation of mathematics teachers, in many European countries that work with a detailed National Curriculum, teachers in a particular area simply can not make their own curriculum (unless a particular set of schools obtain some status as experimental pedagogical testbeds). Similarly, increasing the use of PAT from 25 to 70 days within a single year (as the authors propose) would not be an easy thing to achieve within the constraints of every school timetable (if the children spend much more time on algebra with PAT, on what other subjects do they spend less time?). Finally, it appears that the PUMP project itself required an enormous use of resources — teachers' and researchers' time, as well as funding new computer-labs.

The main message of this section could be reduced to the following. Putting advanced educational technologies into schools is not only a matter of designing systems that are adapted to students' needs, it also requires 'opening up' on several fronts : '*opening up*' the curriculum so that an appropriate role is found for the kind of educational support that computers can provide, '*opening up*' the systems so that different types of actors (teachers, administrators, political deciders, ...) in the educational system can understand how they work, and '*opening up*' the actors themselves to the technology (i.e. *training* them). These are difficult problems to solve ; they need to be addressed at (trans-)national levels. For these reasons, although I think that increasingly 'open' systems will be used increasingly, they are unlikely to become very widespread within the next ten years. If, during the next ten years, computers themselves, together with cheap access to internet, could be available for nearly all school students, using some form of educational software of whatever kind, then that would already be a significant step towards providing the kind of background into which some kinds of intelligent educational systems will find their appropriate place.

It would of course be possible to be 'revolutionary' and claim that all of the above remarks are irrelevant: educational systems will be so transformed in the future to the extent that schools, as institutions as we currently understand them, will no longer exist, students will be elsewhere, learning what they want, how they want, linked at a distance. But I really do not think so ... at least not in the next ten years in Europe.

## 4.2 Other types of educational situations for ITS and other AIED systems

Schools as institutions are not, I think, ready for individualised ITSs in the near future. But these and similar systems may be appropriate for *other* types of educational situations. These situations will, I think, be characterised by a decrease in institutional (and especially teacher) responsibility and an increase in learner autonomy with respect to what is learned, how and when. I think that such situations include learning out of school time, continuing adult education, some types of training in professional contexts and types of "learning for leisure". In the latter case, the principal criteria are personal and operational: that the learners themselves are satisfied with what they have learned, and that they perceive that as being useful and rewarding in some aspect of their lives. If someone wants to learn the Finnish language in their own time, for their own personal reasons, then why wouldn't they want to buy a CD-ROM for that, or a more or less sophisticated ITS program, if it exists? Once the type of learning relates to some type of socially legitimised norm — for example, learning the highway code, learning principles of first-aid — then the design of the AIED system will have to take those norms into account; and that involves addressing those social norms on a national or trans-national level. In fact, many educational CD-ROMS are available today that are closely tied to national curricula. To my knowledge, these are not actually used *in* school lessons, but rather by students in their own time at home, for helping with revision for exams, or homework in general. The reason why this is so clearly does not only relate to teachers' understanding and acceptance of the software, as discussed above, but also to something much more obvious : if



there isn't a computer and the software to go with it for every student in the class (or even one between two) then — some, including myself, would argue — it's simply not fair to provide those resources to some but not all students.

#### **4.3 What other types of computational educational artefacts for classrooms?**

AIED systems do not only have to take social situations into account, but they also can not be developed without heed for ways in which computers are transforming society in general.

In simple (simplistic?) terms, a great deal of educational intervention comes down to providing the information that best suits the learners' needs at the right time, within the constraints of the curriculum, as well as providing the right kind of critical feedback. In the first case, there are two main ways of doing this, each of which has its own difficulties: *generate* or *search*. In the case of "generate", the content of an adapted educational intervention — such as an explanation — is generated dynamically on the basis of a general underlying knowledge representation, and perhaps some model of the student's knowledge and goals. In the case of "search", it is assumed that a vast space of possible presentations of information already exists, and the problem is to find the appropriate content and presentation for the current purpose.

Given the growing amount of information that is available on the World Wide Web, I think — perhaps with most people — that the "search" alternative will become more and more important in the coming years. In that case, AIED research will find its most important application (and it already has to some extent — see e.g. Brusilovsky, Ritter & Schwarz, 1997; Stern, Woolf & Kurose, 1997; Vassileva, 1997) in providing more or less intelligent search engines for the Web. The major problem is, of course, to provide a systematic and internationally standardised representation of the semantics of multimedia information sources. It is important to realise that such engines will be useful for certain types of learning tasks but not others. For example, in problem-solving domains such as branches of elementary mathematics, generation of adapted feedback, provided that it is understandable by and acceptable to the teacher, may well be most appropriate. For project work — e.g. "carry out a project on the religion and art of Ancient Egypt" — the information-search alternative may be more appropriate; and this is a form of task that is already well understood by teachers in existing educational institutions (e.g. teaching students how to use conventional documentary sources, such as libraries).

Other uses of computational models in education may also become more important, providing that they respect the knowledge that the educational actors already have. For example, I can see no reason why a type of student model that provides an abstract summary of what students have already done, and what problems were experienced, should not be a stimulating object of reflection for students and teachers (cf. Bull & Broady, 1997).

Collaborative educational technologies based on computer-mediated communication across the internet are currently in vogue within the AIED community. There again I think that the problems are not just theoretical and technical, but rather relate to identifying educational situations in which these technologies really bring something new and useful. For example, one such situation is learning foreign languages in Europe: the advantages of a student in Naples who needs to learn German being able to discuss with students in Germany are quite obvious. Just as language teachers have always had to establish contacts with schools abroad (e.g. to organise "pen-friends" or visits), they will have to set up and manage distributed learning communities. And managing such communities, so that beneficial interactions do take place, is a lot of work. So where could model-components come in? One possibility would be in identifying groups of learners who could profitably work together (e.g. Hoppe, 1995). Another possibility — and I think, a very important one — would be in providing (intelligent?) support for teachers who have to manage distributed learning communities.

In summary, the crucial problem facing the model-as-component-of-educational-artefact current in AIED research, over the coming years, is to identify the educational situations in which these model components can really be useful, within social and institutional constraints. I think that research will emphasise more and more the support of the teacher's role in varied

situations. The phenomenal increase of information that is available on the web is a fact that can not be ignored. This will lead to the need for adapting existing information sources towards educational ends, and the need for help for students and teachers in finding the information that is most adapted to their needs.

## 5. MODEL AS BASIS FOR DESIGN OF COMPUTATIONAL ARTEFACTS FOR EDUCATION

In his opening address to the "Learning Sciences" conference, at Northwestern University in 1991, Professor Roger Schank raised the following (rhetorical) question, with respect to multimedia training systems developed in his institute: "Where's the AI?". The response that he immediately provided was: "It's in the design". Given Schank's earlier research on memory, explanation and language understanding, it was apparent that what he meant by AI "in" a system was an AI *model* as a system component (e.g. a student model, a natural language understanding module). But how exactly can a model be *a basis for design*? What does it mean to say that a theory or AI model is *in* the design of a computational educational artefact?

In section 3 above, I described what types of entities AIED models of educational processes might be, and in section 4, I described some ways in which these models could become useful components of educational artefacts. But the model-component approach is not, I think, the only one: AIED models *can* be useful tools for the design of educational artefacts, as I will discuss below. A theory or a model of an educational process can be a basis for design of artefacts in a number of different ways.

One possibility is that *a theory gives rise to a design method* (e.g. Clancey, 1993). For example, a theory of learning as a situated activity, a social practice, can then be applied to the research and design activity itself — i.e. design of educational artefacts as a social practice involving computer scientists, content experts, teachers, students, .... In this case, the theory quite consistently does not say exactly what the educational artefacts will be; rather, it says that — in accordance with the underlying theory — the artefacts will aim to enable situated learning; and — in accordance with the situated view of what a research activity is — it says something about how they will be designed. It is an open question as to the extent to which that design process can itself be an object of scientific reflection and ... modelling ... as well as the usefulness of such a formalisation/rationalisation. One use that I can see is in training people to become educational designers. Training by legitimate peripheral participation or cognitive apprenticeship presumes that sufficient real situations exist into which the trainee can become integrated. It is well established with respect to any other profession that such a form of training — e.g. of engineers, who alternate on the job experience with formal training — can not constitute all of the training curriculum. So we do need to reflect upon such a process, and to systematise it to some extent, so that others can be trained to do it, at least as a preparation for entering into the real-life situation itself. In sum: there are good reasons for wanting to systematise situated, or otherwise understood, educational design processes: so that others can learn to become designers.

A second possibility is that a theory or model of learning gives rise quite directly to the range of activities that an educational system does and does not support. For example, a model of learning as reinforcement, chunking and tuning of cognitive procedures leads directly and coherently to the choice of correcting the learner's errors immediately. In fact, this possibility corresponds to the model-component approach described above, which reveals that it is a specific case of the "model as basis for design" approach.

A third possibility involves designing the tools provided for learners in a quite direct relation to theories and models. For example, a theory that views the knowledge to be acquired as embedded in artefacts will clearly lead to providing tools that facilitate such embedding of the learner's activity — the tools keep traces of that activity. A concrete example of use of formal models in design of an educational artefact — without transposition of a computational-cognitive model into a run-time system component — is provided by the C-CHENE system

(Baker & Lund, 1997). The question at issue was to design communication tools that enabled and facilitated collaboration in problem-solving based around a graphical interface (for learning about modelling energy in physics). The system was based on existing models of collaborative problem-solving (Baker, 1994, 1995) and dialogue management (Bunt, 1995) that, together with analysis of previous interaction corpora, directly formed the basis of the communication tools that we provided (e.g. the set of speech act buttons, the mechanisms for coordinating activity). The analytical model described above in section 3.2 (Baker, in press-a) was developed *during* the modelling (design, implementation, experimentation) process itself. This reveals how production of artefacts is rarely a simple 'application' of models, but also a pragmatic means of elaborating them.

I think that it is possible and desirable to combine all three of the above possibilities in using theories and models as the bases for design of educational artefacts. This would involve both working with social actors in specific educational situations, and proposing new possible tools on the basis of existing formal models. Whilst potential users of these artefacts often have firm and clear beliefs as to what their needs are, they are not always aware of technological possibilities. It is not just a matter of finding out what they want, but also of proposing new possibilities that could transform the social actors' activities in previously unthought-of ways. For example, we have been able to use C-CHENE as a basis for a new form of teacher training (Lund & Baker, 1999). We could not have developed that by working initially with teachers, since they would never have seen the need for such a tool. Rather, once we had developed it, on the basis of existing models of collaboration and dialogue, the teachers could see a creative extension of what their teaching practice might become (cf. Schön, 1983) using the tool. Further work with teachers is being carried out so that the tool can be appropriated within teaching practice.

In summary, theories and models of educational processes, as they are elaborated within AIED research, can constitute the bases for design of educational artefacts in a number of different ways. Using a computational model of some human capacity as a system component is only one of those ways; others involve — quite legitimately — a more or less direct relationship between theories and models and design processes and products.

Finally, one may often be tempted to ask the following: "But couldn't the computer-based learning environment have been developed *without* that theory or model? Isn't the theory just a verbal gloss on the system, an *a posteriori* justification or rationalisation?". One possible answer would be: "yes, perhaps it *could* have been designed without that theory, who knows?; but as it turns out, it *wasn't*". To me, this points to the necessity for researchers in our field to make as explicit as possible the relations between theory, model and artefact, this being essential for understanding how to take research forward. Whilst it is not, in my opinion, legitimate to require a 'direct' relation between theory, model and artefact design, the nature of those relations is always a legitimate object of scientific debate. In my view, one of the defining characteristics of AIED research is precisely the fact that it aims for theoretically-founded design of computational educational artefacts, in conjunction with the elaboration of models of educational processes.

## 6. SYNTHESIS AND CONCLUDING REMARKS

I have sketched a personal and prospective view of AIED research that turns on three possible roles for models: as *scientific tools*, as *components* of computational educational artefacts, and as *bases for design* of such artefacts.

In terms of the first role, my view is that AIED research, over the past three decades, has already mapped out a vast space of phenomena to be studied. We do not need to extend the space of phenomena, but rather to extend the range of theoretical tools from those available in cognitive science, and to adopt a wider (yet more strict) notion of what is and what is not a model. Specifically, and in terms of how I defined models themselves, I claimed that there is no *a priori* reason why interesting models should not be developed, that extend the notion of

'cognition' to embrace action and perception, as embedded in artefacts and social relations. AIED research should and will, I think, open out to a greater extent than is currently the case, into cognitive science, considered in the widest sense of the term. The role of a model, as scientific tool, is to help us to explain, to develop theory, and to predict. As such, any model abstracts from reality. Failure to take a particular phenomenon into account does not invalidate a model, it just restricts its usefulness.

In terms of the second role — models as components — I claimed that individualising ITS are not currently adapted to existing educational practices, largely because of, on a micro-level, problems associated with failing to take teachers, and other social actors, into account. Either we must adapt the components and the artefacts, or else change educational systems; and no doubt, most researchers aim for some realistic combination of both. Depending on the culture concerned, there may be a greater or lesser difference between the timescales of institutional and technological change. I proposed that ITS will, in the near future, be most appropriate for social situations that are less norm-based than most state education systems. Within such educational situations, intelligent information search for learners using the Web, rather than intelligent explanation generation, will come to the forefront in the near future, depending on the type of learning task involved. Intelligent explanation generation, and help systems in general, may turn out to be more important for *teachers* rather than for learners, in, for example, distributed learning communities. Models as intelligent components of educational artefacts have, I think, an important role to play in the near future; it is simply that their uses may not be in the situations that AIED researchers originally thought.

Finally, once we remember that (of course) models are not, by their nature, necessarily computational, this opens up a wide range of possible ways in which theories and models can form the bases of design of educational artefacts. What is required is that the specific nature of the relations between theory, model and design of artefacts be made as explicit as possible<sup>11</sup>, as legitimate objects of scientific discussion and as means of generalising findings towards re-design. Personally, I believe that theories and models will find their most effective application in design of collaborative distributed educational technologies.

I conclude with some brief remarks on the unity and future of AIED research, as a field. Given all the possible evolutions of AIED research that I have sketched, isn't there a strong possibility that AIED could dissipate into educational research and/or that part of cognitive science that is concerned with learning and teaching? Perhaps, and after all, why not? But I do not think so, and for the following reasons. In terms of the particular view of AIED research I have outlined above, what makes piece of research AIED research is, quite simply, that it has something innovative to say about *all three* of the possible roles of models, with a greater or lesser emphasis being put on each. Concretely, this means that the research in question proposes a specific, explicit and coherent set of relations between: (1) a theory, (2) a model, (3) an experimental field of educational phenomena, (4) computational-educational artefacts, whose use is part of (3), and (5) an educational design process. It is not enough to propose a model of an educational phenomenon; the research must also describe how the model relates to theory, how it is relevant to study or design of artefacts for teaching and learning, and how that design might proceed. This means that AIED research is very complex, and very difficult to carry out.

I think that those constraints will continue to be sufficient for distinguishing a specific field or area of research, whether it is called AIED or something else.

## Acknowledgements

This work was financially supported by the CNRS. Collaborative work with a number of colleagues over a number of years has helped and forced me to make my views on AI and Education research methodology more explicit. I would particularly like to thank (in

---

<sup>11</sup> John Cook's recent work (Cook, 1998a, 1998b) describes precisely such a principled relation between theory, model, corpus data and artefact for the case of an ITS for learning musical composition.

alphabetical order) Paul Brna, John Cook, Pierre Dillenbourg, Tia Hansen, Kristine Lund, Matthieu Quignard and Andrée Tiberghien. They do of course bear no responsibility for any errors or short-sightedness in the views I have expressed here. Finally, I would also like to thank three anonymous reviewers for helping me to make some of my arguments more precise and some of my claims less excessive.

## References

- Allwood, J., Nivre, J. & Ahlsén, E. (1991). *On the Semantics and Pragmatics of Linguistic Feedback*. Gothenburg Papers in Theoretical Linguistics No. 64. University of Gothenburg, Department of Linguistics, Sweden.
- Baker, M.J. & Bielaczyc, K. (1995). Missed opportunities for learning in collaborative problem-solving interactions. *Proceedings of AI-ED '95 : World Conference on Artificial Intelligence in Education*, Washington D.C., August 1995. Washington: AACE.
- Baker, M.J. & Lund, K. (1997). Promoting reflective interactions in a CSCL environment. *Journal of Computer-Assisted Learning*, 13, 175-193.
- Baker, M.J. (1994). A Model for Negotiation in Teaching-Learning Dialogues. *Journal of Artificial Intelligence in Education*, 5(2), 199-254.
- Baker, M.J. (1995). Negotiation in Collaborative Problem-Solving Dialogues. *Dialogue and Instruction: Modeling Interaction in Intelligent Tutoring Systems*, (eds.) R.-J. Beun, M.J. Baker & M. Reiner, pp. 39-55. Berlin: Springer-Verlag.
- Baker, M.J. (1996). Argumentation et co-construction des connaissances. *Interaction et Cognitions* 2 (3), 157-191.
- Baker, M.J. (in press-a). Forms of cooperation in dyadic problem-solving. To appear in T. H. Benchekroun & P. Salembier (Eds.) *Cooperation and Complexity*. Berlin: Springer-Verlag.
- Baker, M.J. (in press-b). Argumentation and Constructive Interaction. In J. Andriessen & P. Coirier, (Eds.), *Foundations of Argumentative Text Processing*. Amsterdam : University of Amsterdam Press.
- Baker, M.J., Hansen, T., Joiner, R. & Traum, D. (1998). "Grounding" for Intersubjectivity and Learning. Paper presented at the *Fourth Congress of the International Society for Cultural Research and Activity Theory* (ISCRAT 98), University of Aarhus, Denmark, June 7-11, 1998. Available as COAST Research Team technical report at the following address: <http://signserver.univ-lyon2.fr/SIR/GRIC-COAST/coast.reports.html>
- Baker, M.J., Hansen, T., Joiner, R. & Traum, D. (1999). The role of grounding in collaborative learning tasks. In P. Dillenbourg (Ed.) *Collaborative Learning: Cognitive and Computational Approaches*, pp. 31-63. Amsterdam: Pergamon/Elsevier Science.
- Barth, E.M. & Krabbe, E.C.W. (1982). *From Axiom to Dialogue: A philosophical study of logics and argumentation*. Berlin: Walter de Gruyter.
- Brusilovsky, P., Ritter, S. & Schwarz, E. (1997). Distributed Intelligent Tutoring on the Web. *Artificial Intelligence in Education: Knowledge and Media in Learning Systems* [Proceedings of AIED-97, Kobe Japan], pp. 482-489. Amsterdam: IOS Press.
- Bull, S. & Broady, E. (1997). Spontaneous Peer Tutoring from Sharing Student Models. *Artificial Intelligence in Education: Knowledge and Media in Learning Systems* [Proceedings of AIED-97, Kobe Japan], pp. 143-150. Amsterdam: IOS Press.
- Bunt, H. C. (1989). Information dialogues as communicative action in relation to partner modelling and information processing. In Taylor, M. M, Néel F, & Bouwhuis D. G. (Eds) *The structure of multimodal dialogue*, 47-74. North-Holland: Elsevier Sciences Publishers.
- Bunt, H. C. (1995). Dialogue Control Functions and Interaction Design. In Beun, R. J., Baker, M.J. & Reiner M. (Eds.), *Dialogue and Instruction, Modeling Interaction in Intelligent Tutoring Systems*. Proceedings of the NATO Advanced Research Workshop on Natural Dialogue and Interactive Student Modeling, 197-214. Berlin: Springer Verlag.
- Clancey, W.J. (1993). Guidon-Manage Revisited: A Socio-Technical Systems Approach. *Journal of Artificial Intelligence in Education*, 4(1), 5-34.

- Clark, H. H & Brennan, S. (1991). Grounding in communication. In L.B. Resnick, J.M. Levine & S.D. Teasley (Eds.) *Perspectives on Socially Shared Cognition*, pp. 127-149. Washington DC: American Psychological Association.
- Clark, H. H., & Schaefer, E. F. (1989), Contributing to discourse. *Cognitive Science*, 13:259-294.
- Cook, J. (1998a). Mentoring, Metacognition and Music: Interaction Analyses and Implications for Intelligent Learning Environments. *International Journal of Artificial Intelligence in Education*, 9, 45 - 87. [<http://cbl.leeds.ac.uk/ijaied/>]
- Cook, J. (1998b). *Knowledge Mentoring as a Framework for Designing Computer-Based Agents for Supporting Musical Composition Learning*. Unpublished PhD thesis, Computing Department, The Open University (GB).
- Dillenbourg, P. & Self, J. (1992). A computational approach to socially distributed cognition. *European Journal of Psychology of Education*, 7 (4), 353-372.
- Dillenbourg, P. & Traum, D. (1996). Grounding in multi-modal task-oriented collaboration. In P. Brna, A. Paiva & J. Self (Eds.), *Proceedings of the European Conference on Artificial Intelligence in Education*, pp. 415-425. Lisbon, Portugal, September 1996. Lisbon: Edições Colibri.
- Dillenbourg, P. (1991). *Human-Computer Collaborative Learning*. PhD thesis (unpublished), University of Lancaster (UK), Department of Computing.
- Dillenbourg, P. (1999). What do you mean by "Collaborative Learning"? In P. Dillenbourg (Ed.) *Collaborative Learning: Cognitive and Computational Approaches*, pp. 1-19. Amsterdam: Pergamon/Elsevier Science.
- Dimitrova, V., Self, J. & Brna, P. (1999). The Interactive Maintenance of Open Learner Models. S.P. Lajoie & M. Vivet (Eds.) *Artificial Intelligence and Education, Open Learning Environments: New Computational Technologies to Support Learning, Exploration and Collaboration* [Proceedings of AI-ED 99, Le Mans], 405 - 412. Amsterdam: IOS Press.
- Edmondson, W. (1981). *Spoken Discourse : A model for analysis*. London: Longman.
- Greeno, J. (1989). Situations, mental models and generative knowledge. In Klahr, D. and Votovsky, K. *Complex Information Processing: the Impact of H.A.Simon*. pp. 285 - 308. Hillsdale, NJ: Lawrence Erlbaum.
- Hoppe, H. U. (1995). Using multiple student modeling to parametrize group learning. In J. Greer (Ed.), *Proceedings of the Seventh World Conference on Artificial Intelligence in Education* (pp. 234-241). Charlottesville, VA: Association for the Advancement of Computing in Education.
- Hoppe, H.U., Baloian, N. & Zhao, J. (1993). Computer support for teacher-centered classroom interaction. In *Proceedings of the International Conference on Computers in Education*, 4, 459-475.
- Inaba, A. & Okamoto, T. (1997). Negotiation Process Model for Intelligent Discussion Coordinating System on CSCL Environment. *Artificial Intelligence in Education: Knowledge and Media in Learning Systems* [Proceedings of AIED-97, Kobe Japan], pp. 175-182. Amsterdam: IOS Press.
- Koedinger, K.R., Anderson, J.R., Hadley, W.H. & Mark, M.A. (1997). Intelligent Tutoring Goes to School in the Big City. *International Journal of Artificial Intelligence in Education*, 9, 45-87. [<http://cbl.leeds.ac.uk/ijaied/>]
- Kuhn, T.S. (1962). *The Structure of Scientific Revolutions*. International Encyclopedia of Unified Science, Volume 2, Number 2. London: University of Chicago Press.
- Leontjev, A. N (1981). The problem of activity in psychology. In J. V. Werstch (Ed.) *The concept of activity in soviet psychology*, pp. 37-71. Armonk, NY: Sharp.
- Lund, K. & Baker, M.J. (1999). Teachers' collaborative interpretations of students' computer-mediated collaborative problem-solving. S.P. Lajoie & M. Vivet (Eds.) *Artificial Intelligence and Education, Open Learning Environments: New Computational Technologies to Support Learning, Exploration and Collaboration* [Proceedings of AI-ED 99, Le Mans], 147 - 154. Amsterdam: IOS Press.

- Lyotard, J.-F. (1979). *La Condition Postmoderne: Rapport sur le savoir*. Paris: Minuit.
- Mammen, J. (1993). The elements of psychology. In N. Engelsted, M. Hedegaard, B. Karpatschof & A. Mortensen (Eds.) *The Societal Subject*, pp. 29-44. Århus: Århus University Press.
- Perret-Clermont, A.-N., Perret, J.-F. & Bell, N. (1991). The social construction of meaning and cognitive activity in elementary school children. In L. Resnick, J. Levine & S. Teasley (Eds.) *Perspectives on Socially Shared Cognition*, pp. 41-62. Hyattsville, MD: American Psychological Association.
- Roschelle, J. & Teasley, S.D. (1995). The construction of shared knowledge in collaborative problem solving. In C. O'Malley (ed) *Computer Supported Collaborative Learning*. NATO ASI Series, pp. 69- 97. Berlin: Springer-Verlag.
- Schön, D. (1983). *The Reflective Practitioner*. Aldershot (GB): Arena Ashgate.
- Self, J.A. (1996). *Computational mathematics: towards a science of computer based learning system design*. Technical Report 96/23, Computer Based Learning Unit, University of Leeds.
- Stern, M., Woolf, B.P. & Kurose, J.F. (1997). Intelligence on the Web? *Artificial Intelligence in Education: Knowledge and Media in Learning Systems* [Proceedings of AIED-97, Kobe Japan], pp. 490-497. Amsterdam: IOS Press.
- Tiberghien, A. (1994). Modeling as a basis for analyzing teaching-learning situations. *Learning and Instruction* 4, 71-87.
- Tiberghien, A. (1996). Construction of prototypical situations in teaching the concept of energy. G. Welford, J. Osborne & P. Scott (Eds.) *Research in Science Education in Europe*, pp. 100-114. London: Falmer Press.
- VanLehn, K., Jones, R.M. & Chi, M.T.H. (1992). A model of the self-explanation effect. *Journal of the Learning Sciences*, 2, 1-59.
- Vassileva, J. (1997). Dynamic Course Generation on the WWW. *Artificial Intelligence in Education: Knowledge and Media in Learning Systems* [Proceedings of AIED-97, Kobe Japan], pp. 498-505. Amsterdam: IOS Press.
- Wertsch, J. V. (1991): *Voices of the Mind. A Sociocultural Approach to Mediated Action*. USA: Harvester Wheatsheaf.
- Wittgenstein, L. (1980). *Culture and Value* [Vermischte Bemerkungen]. G.H. von Wright & H. Nyman (Eds.), trans. P. Winch. Oxford: Basil Blackwell.