Designing Human-Computer Collaborative Learning

Article · November 1994

DOI: 10.1007/978-3-642-85098-1_13 · Source: CiteSeer

Designing Human-Computer Collaborative Learning

Pierre Dillenbourg and John Self

1994

AAI/AI-ED Technical Report No.91

(to appear in C. O'Malley (ed.), *Computer -Supported Collaborative Learning*, Berlin:Springer-Verlag)

Computing Department
Lancaster University
Lancaster
LA1 4YR
UK

aai@comp.lancs.ac.uk
http://www.lancs.ac.uk/computing/research/aai-aied/
ftp.comp.lancs.ac.uk (148.88.8.9) /pub/ai/

Designing Human-Computer Collaborative Learning

P. DILLENBOURG & J. SELF
Computing Department, Lancaster University, U.K.

1. Introduction

Most chapters of this book investigate how the collaboration between human learners may be supported or modelled by a computer. Our approach is quite different. We envisage here the collaboration between a human learner and an artificial learner, simulated by the computer: the student and the computer learn together by using the same micro-world. We will refer to the computerized learner as the "co-learner". The main structure of such a system is shown in figure 1.

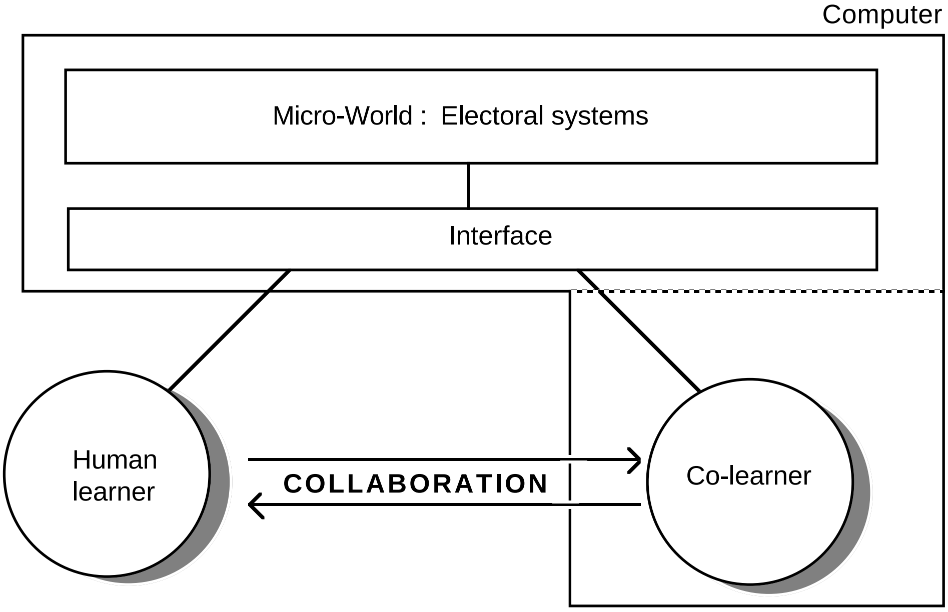

Figure 1: Main structure of PEOPLE POWER

The domain chosen for this system (called "PEOPLE POWER") belongs to politics: the goal is that learners discover the biases present in the various European electoral systems. An electoral **bias** is defined as a difference between the people's preferences distribution in the country (before the elections) and the distribution of seats in the parliament (after the elections). This definition of a bias reveals that we have focussed on the search for a *proportional* system rather than the undefinable *democratic* system. Such bias must be related, through experiments, to the various characteristics of electoral systems such as the size of constituencies, the number of seats per constituency, the way of expressing preferences by a vote, the way of counting the votes and the decision rules allocating seats with respect to votes. Here is an example of bias we expect the learners will discover: *"Within a system based on absolute majority and on single-place constituencies , a party can theoretically win the elections (i.e. get 51% of the seats) with 26% of the votes (51% of the votes in 51% of the constituencies) . If the system is based on relative majority this percentage may be even less."*

Discovering such bias involves building a deeper understanding of electoral mechanisms. Every adult has some naïve model of how elections are performed. Few are conscious that electoral results may be related more to intrinsic characteristics of an electoral system than to the people's actual preferences. This discovery process will be performed by running experiments in order to compare variations of electoral results obtained with the same preferences but with different systems. The role of the micro-world is to make easier the elaboration and operation of an electoral experiment. The role of the co-learner is to share the difficult task of designing interesting experiments and understanding their results.

2. Rationale for the human-computer collaborative learning.

This approach is not to be understood as an investigation of the benefits we might expect from replacing a human learner by a computerized one. This position would be indefensible. Instead, the rationale of our approach must be understood within the field of Intelligent Tutoring Systems (ITS) and Learning Environments, where the debate about student control vs computer control has always been with us. After the over-optimism of LOGO enthusiasts, most researchers have submitted to the need to offer stronger guidance to help students. We view this position as defeatist and conservative. If, after several years of passive schooling, students do not show enough autonomy in learning-alone settings, this does not mean that autonomy may not be developed. Offering stronger guidance guarantees short-term effectiveness, but in a longer perspective just reproduces the schooling effects.

Beside the increase of guidance, an alternative solution is to develop systems which give the students the opportunity of (re-)developing the ability to learn on their own. The main focus of such systems would be the learner's metacognitive skills. Metacognition is a difficult object of investigation since its effects are difficult to measure and to control. We think it is time to risk an escape from easy-to-observe educational issues if we want to have a real impact on the educational system.

We also like to view our system as an attempt to explore architectures which radically differ from the traditional ITS structure: it has no domain model, but tries to acquire it by interacting with the student; it has no tutoring component, the co-learner having no hidden didactic plans.

At its current stage of development, this human-computer collaborative approach raises more questions than it brings answers. Most of these questions are fundamental: the mechanisms of collaboration, the interaction between a human and a computer, the plausibility of machine learning techniques. We do not expect to bring complete answers. The goal of this chapter is to describe the conceptual framework which - we believe - relates our main goals expressed in terms of student autonomy and the implementation of a collaborative learning system. This relationship constrained the design process of PEOPLE POWER.

3. Designing collaboration

Peer learning or peer problem solving covers a very large range of collaborative styles, some of them having more benefits than others. The kind of collaborative style depends on many factors: the learners' characteristics, their relationship, the nature of the task or the context. For instance, the distribution of the roles among the partners or the domination of one learner over the other may inhibit the emergence of conflicts (Joiner, 1989). This domination is especially negative when learning requires that both learners perform concrete manipulations. Salomon and Globerson (1989) identify a list of factors which reduce peer learning efficiency: the "free rider effect", the "sucker effect", ...

We may partially determine the collaborative style by carefully modifying some factors. In our case, we can act on two factors: the nature of the task and the behaviour of the co-learner. Our design choices will depend on the kind of collaborative style we have in mind. This collaborative style is itself a function of the kind of educational benefits we hope students will gain from collaboration. Consequently, we will first describe our educational goals in building PEOPLE POWER, and then try to deduce some desirable characteristics for the learning task and the co-learner behaviour. Our approach is represented by the deductive process showed in figure 2.

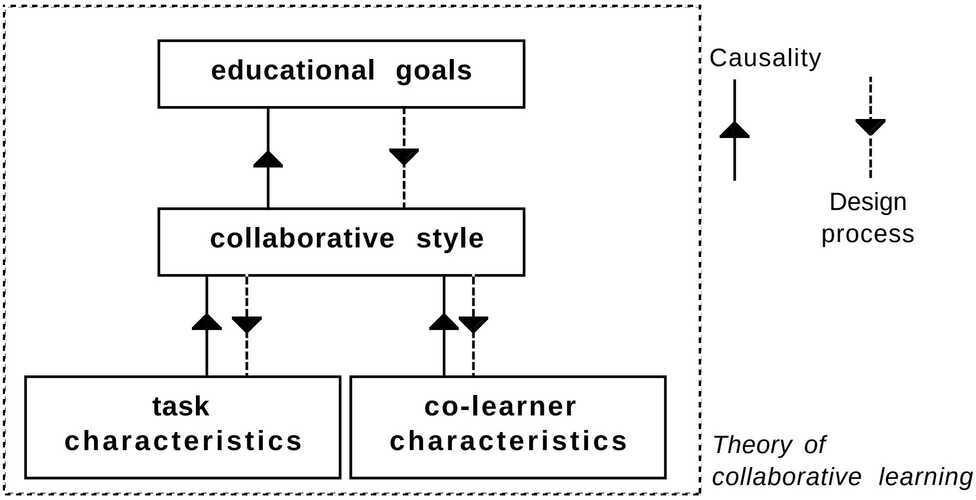

Figure 2: Components of the design process

However, as is shown in this book, it is obvious that there exists no complete and consistent theory available for safely guiding this deductive process (Saunders, 1989). Hence, we do not claim we will suppress the empirical tuning stage. Nevertheless, if we want to learn from the system experimentation, we need to describe the hypotheses underlying this deductive process and to present some elements which partially support them (section 4). Our system may be seen as an opportunity to find complementary evidence.

4. Educational goals: focus on metacognition

Our approach to collaborative learning is centred on the metacognitive benefits of collaboration. As stated in section 2, we see metacognitive skills as a way of increasing the student's ability to learn alone. Metacognition refers to the set of the subject's cognitive activities which have as object her own cognitive process. Two kinds of metacognition are generally discriminated (Nguyen-Xuan, 1989):

- metacognitive knowledge: the subject's naïve theories about intellectual processes (e.g. we remember better things we have written), which corresponds to the AI concept of *metaknowledge* ;

- "pure" metacognition: some self-awareness of one's own knowledge, which corresponds to the AI concept of *computational reflection*.

The "meta" prefix has been besmirched in cognitive science during the last decade, but the educational importance of metacognitive processes convinces us that it is still worth concentrating research efforts on this topic. However, this does not mean that we share a hyper-rational view of a human learner, using metaknowledge for building plans for each of her actions. Our view is closer to the idea of **situated actions**: "*every course of action depends in essential ways upon its material and social circumstances*" (Suchman, 1987, p.50). This approach probably reduces the importance of some eventual (rare) domain-independent metaknowledge, but it emphasizes the role of metacognition and (acquired) domain-specific metaknowledge for performing "on-the-fly" intelligent adaptation to the circumstances of action: "*...plans are a constituent of practical action, but they are constituent as an artifact of our <u>reasoning about</u> action, not as the generative <u>mechanism of</u> action*" (ibid, p.39, original emphasis).

The design of our system is based on the relationship between our metacognitive goals and the collaborative setting. We describe three mechanisms by which collaboration may promote metacognitive skills. Albeit we present some evidence supporting these mechanisms, they still have in our eyes the status of hypotheses.

<u>4.1 Mechanism #1: Verbalizing strategic decisions.</u>

Our first hypothesis is that, if learners have to agree on the learning or problem solving strategy, they are required to <u>make explicit strategic decisions</u>. In other words, if strategic decisions are necessary (i.e. if it is a multiple steps learning task), they will be the object of verbal agreement like any other decision made between the two learners.

Some evidence for this hypothesis is found in the work of Balacheff (1989). He submitted the following problem to 13-14 year old students: "*Give a way of calculating the number of diagonals of a polygon once the number of vertices is known*". Here are several examples chosen among the protocols he collected. These examples of interaction explicitly concerned strategic decisions (we translated the original French protocols):

- *"... we should first find out what a polygon is."*

- *"We should try with 4 vertices, it would be easier."*

a- *"So, what do you want to do ?"*
b - *"Me ? Thinking from our two developed theories; maybe to group them and develop them rather than starting various trials which disturb us"*
a- *"OK, but I think it will lead us nowhere."*
b- *"Maybe it will lead us nowhere, but it will lead us further on than your trials (he laughs), because your trials did not lead us further on"*

a- *"Usually a theory may be applied to every case."*
b- *"In grammar, there are exceptions."*
a - *"Yes, but not in mathematics."*
b- *"Yes, there are exceptions in mathematics."*
a- *"Yes, but fewer ..."*

These strategic decisions enter into the metacognitive zone when they are made explicit and communicated in order to reason on past or future actions. Such reasoning is precisely required by the negotiation involved when the learners search to agree on decisions. Once these very context-specific decisions have been informally assessed through local experiences, they may be progressively de-contextualized and turned into less context-specific metaknowledge. Inversely, some of these decisions may result from attempts to apply general metaknowledge to the context and may be progressively refined (through further experiences) into efficient but specific heuristics.

4.2. Mechanism #2: Acquisition of reflective skills

If our goal is to develop metacognition, we may wonder how one may become aware of one' own knowledge. Is it a permanent connection between levels of knowledge or some automatic maintenance process? Or is it instead the result of some intentional activity or at least of some trained mental habit? This last hypothesis is the one we adopt: bringing into consciousness some implicit piece of knowledge Ki does not occur by itself, but results from some reasoning which enables access of Ki from some explicit piece of knowledge Ke. We will use the terms "implicit" and "explicit" instead of "conscious" and "unconscious" because the latter have too many connotations in psychology.

The relationship between Ki and Ke may be any kind of association. Nevertheless, our educational context leads us to concentrate on the logical relationships. The intentional activity so described corresponds to the definition Dewey (1933) gave to the word **reflection**: "*reflection is an active, persistent and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it*" .

Collaboration promotes reflection because the scepticism or the criticism of one learner constrains the other learner to justify or to defend himself. "*Reflection is most profound when it is done with the aware attention of another person*" (Knights, 85). We may rewrite this hypothesis in the formalism proposed by Elsom-Cook in his contribution to this book:

```
aware (A1, Ki)    <-  ext-conflict (A1,A2,Ke,Kx)    &
                      want (A1, believes (A2,Ke))   &
                      believes (A1, Ke <- Ki)
```

The meaning of ext-conflict includes not only criticism resulting from clear conflicts (which are seldom - see Blaye's chapter), but any discrepancy of knowledge between an explainer and an explainee. Reflective activities might explain the benefits gained by the more knowledgeable member of the peer group when she gives an explanation to the co-learner (Webb, 89): the explainer may discover the importance of the `Ke <- Ki` relationship, previously implicit, and hence perceive more clearly the causal structure of the learning object.

We found some evidence for this hypothesis in the studies of Chi et al. (1989) on self-explanation. In these studies, students learned individually to solve physics problems. The learning material included some theory and worked-out examples. The authors analysed the explanations that the learners spontaneously produced when looking at the examples. For these authors, self-explanation enables the building of some inference rules, an intermediate step in the proceduralisation of declarative knowledge. These inference rules are more context-specific than the declarative knowledge, because they are generalised from a small set of examples. More precisely, they may emphasize elements which were not noticed by the learner when reading the declarative knowledge, but which appeared as important while explaining examples.

These authors found that the students who produce more numerous and more elaborated self-explanations are also those who more frequently realize their misunderstanding. They hypothesize that the latter characteristic ("purely" metacognitive) might be the cause of the former. The inverse causal relationship is also plausible, especially in a collaborative setting where the explanation is initiated by the other learner's question.

4.3. Mechanism #3: Acquisition of a better model of learning

After several years in schools, students acquire some idea of what learning is: "*Students tend to view learning as a passive experience in which one absorbs knowledge or copies fact into memory. Little of what they do in schools leads them to question that perspective.*" (Lochhead, 1985). It is clear that students with such a passive model of learning may not benefit from unconstrained learning environments. Giving to the students an opportunity to develop a better view of learning is one of the most important goals of education.

We are indeed afraid that current ITSs do not present a very good model of learning: the fact that knowledge may be recorded inside the tutor "domain model" may reinforce the "copy" model of learning (one computer's memory may be copied by another computer). In some ways, a domain model in an ITS is in contradiction with a discovery approach: on one hand, we tell the student that he must acquire knowledge by himself, but on the other hand, this knowledge exists somewhere in an inspectable form. Our students are indeed used to this kind of dissonance. Since they arrived at school, they have intuitively acquired such rules of the learning-teaching game. Nevertheless, it might be worth investigating how ITSs might avoid giving a model contradictory to the behaviour they want to install.

We view collaboration as a tool which might help students to acquire an active model of learning. Our approach is based on Vygostky's claims on the **internalization** of higher psychological functions: "*Every function in the child's development appears twice: first, on the social level, and later on the individual level; first, between people (inter-psychological) and then inside the child (intra-psychological)*" (Vygostky, 1978).

The learning process involving two individuals is similar to the learning process occurring inside an individual: knowledge change results from the confrontation of (partially) incompatible viewpoints. One difference is that the confrontation is harder when the viewpoints belong to different agents, because it is increased by the social pressure and because it is made more explicit through communication. This partially explains why peer interaction promotes conceptual change. The second, and more interesting, difference is that, in collaborative situations, communication makes the learning process observable. In other words, learning as an intra-individual process can be made observable by transposing it as an inter-individual process.

A pair engaged in a collaborative learning task is a micro-society where one can observe learning. We called this micro-society a "**constructorium**", i.e. a laboratory where one can observe the construction of knowledge. Psychologists already use collaboration as a constructorium when they observe peer interaction for collecting thinking aloud protocols (see chapter X,Y and Z of this book). The educational interest is that, inside this constructorium, learners are also observers: they do not simply perform a domain-specific learning task, they attend to some dissection of the learning mechanisms, they can view the dialectical aspect of learning and consequently modify their representation of the learning process.

In other words, our third hypothesis is that the collaborative setting may be used as a constructorium where the learners will observe and internalize a better model of learning. We found some evidence for this hypothesis in two sets of experiments which have used a collaborative setting for promoting a shift **from a passive to an active model of intellectual processes**.

Palincsar and Brown (1984) used **reciprocal teaching** for installing an active model of **reading** by students who have reading difficulties. In reciprocal teaching, the students alternate the role of reader and leader for each section of a text. The leader's role is to constrain the reader to be active during her reading by asking her a question, requiring a summary, a prediction or a clarification. Five of their six subjects reached within 12 days a level similar to average readers at an individual test of text understanding and maintained this level in follow-up measurements (8 weeks later). This experiment may be criticised for the smallness of the sample but similar results have been found through successive analogous experiments. It is interesting to notice that Miyake (1986) has observed a spontaneous distribution of the roles within the pair, similar to the distribution imposed in reciprocal teaching: one learner does the concrete manipulations while the other has a status of criticiser.

Schoenfeld (1987) described the classroom techniques he used for installing a more active model of **problem solving** among his students. Bloom and Broder (1950, quoted by Lochhead, 1985) have emphasised the importance of this model: "*The major difference between the successful and the non-successful problem solvers in their extent of thought about the problem was in the degree to which their approach to the problem might be passive or active*". A passive problem solver typically reads the problem, chooses one way to solve it and keeps trying this way even if it fails. A more active solving model involves frequent problem re-analysis and backtracking to alternative solution paths. Schoenfeld reduced the number of passive solvers among his students, from 60% to 20%, by applying a programme centred on metacognitive issues. About half of his programme was formed by collaborative-like activities: criticising other students working problems (shown on a videotape) and small group problem solving (3-4 students).

These two experiments seem to confirm the ability to use a collaborative setting to promote an active model of some intellectual process, with the restriction that in both experiments, the human teacher kept an important role during some initial phase.

5. Collaborative style

The learners must have equal status, i.e. they must have access to the same control and the same knowledge, and have the same role in the system. We examine first the control issue.

The form of collaboration will not involve any distribution of the work. The learners will design together the electoral simulations. In theory, we should constrain them to find an agreement. Nevertheless, our system will be slightly asymmetric, because our main interest still concerns the real learner's activities rather than the co-learner's. Hence, the co-learner will describe its suggestions relative to the next electoral experiment and the real learner will be free to take them into account for designing the next experiment. However, if the real learner never satisfies the co-learner requests, then the co-learner will be allowed to manifest some dissatisfaction. This slight degree of asymmetry is obviously one of the elements to be empirically tuned.

The computerized co-learner may not have access to hidden elements of knowledge which enable him to cheat during learning. This is a problem because of the importance of background knowledge in machine learning approaches: the co-learner needs some initial knowledge in order to acquire further knowledge. Actually, the real learner will also start this learning session with some knowledge about electoral systems. The solution is to give to the co-learner some initial naïve model of electoral systems and to make this model inspectable by the real learner.

The knowledge used by the micro-world is quite different from that to be acquired by the co-learner. It is compiled knowledge which enables it to run elections (count votes in some area, etc), but does not enable it to reason about elections in order to explain some particular results. However, it is important that the real learner perceives this distinction of knowledge sources and does not consider the co-learner as an expert in electoral issues: it is not possible to have a real discussion with someone who is supposed to be always right.

This distinction between the co-learner's knowledge and the micro-world knowledge is an aspect of the ambiguity inherent to an approach where the computer has to play two different roles: as a micro-world and as a co-learner. This means that the real learner must be able to view the **computer at the same time as a tool for performing operations (micro-world) and as a partner engaged in interaction (co-learner)**. Suchman (1987) claims that one spontaneously favours the second view because of the computer's opacity: "*the personification of the machine is reinforced by the ways in which its inner workings are a mystery, and its behaviour at times surprises us* " (p.16) . This has important implications for the design of the micro-world interface: it has to be as transparent as possible. But it also raises the (unanswered?) issue of knowing if learners are able to adopt different views of the same device, according to the function it assumes at each stage of the interaction.

This issue is a key factor for implementing collaboration between real and computerized learners. The quality of this collaboration will be strongly dependent on how the real learner perceives the co-learner. There is for instance a risk that students attribute didactic intentions to the co-learner. The authenticity of our system requires that the co-learner has no hidden didactic plans. This for instance means that it will ask questions when it really needs the answered information, and not for checking the learner's progress. However, the way students infer the computer reasoning mechanisms seems to be less related to the actual mechanisms than to the students’ a priori expectations towards the computer. This has been exemplified by Weizenbaum's ELIZA program and by Carfinkel's counselor system, whose reasoning was restricted to randomly producing "yes" or "no" answers, but that the users still perceived as driven by some intentional reasoning mechanism (Suchman, 1987)!

This emphasizes the importance of the description of the learning situation which will be given to the real learner at the beginning of the session. It does not mean that we will implement an ELIZA-like co-learner: we see ITS design as a way to investigate cognitive issues (the co-learner design), not to avoid them! The dialogue facilities have to include aspects of interaction which constitute the specificity of collaborative learning: sharing an initial set of meanings, explicit agreement-disagreement messages, explicit requests for explanations and mechanisms managing the dialogue turn taking.

The issue of sharing an initial set of meanings is unsolvable because each term's definition refers to other terms so that a deep analysis rapidly concerns an infinite set of world or common sense knowledge. Like Suchman (1987), we view this shared set of meaning as the result of interaction instead of its condition: divergent meanings produce disagreement or requests for disambiguation which progressively lead to shared meanings. This interactive building of shared meanings is performed through "clarifying subdialogues" (Suchman, 1987), a frequent structure in dialogue analysis. In other words, we think dialogue facilities required for obtaining a shared set of meanings are assumed by the three other characteristics we quoted: (dis)agreement, explanations, turn taking.

We will focus on the co-learner's mechanisms underlying these dialogue facilities instead of on the form of these facilities. This means for instance that our interest concerns more the benefits of giving explanations than the form of the explanations (natural language, telegraph-like sentences, menu sequences,...). We are aware that these aspects of the interface are not simple details and that they may constitute critical aspects of the dialogue, playing an important role in learner's collaborative involvement. It is hence very probable than our current design choices will have later to be revised in order to pay more attention to the concrete aspects of communication.

The learners will communicate via their **notebook**. The first part of the notebook will systematically record the data collected during the experiments, as in REFRACT (Reimann, 1988) or SMITHTOWN (described in Wenger, 1987). The learners are allowed to perform simple data manipulations such as those available in a spreadsheet. The notebook may be compared to the reification tools described by Collins and Brown (1988), which help the learner to reflect upon her experience. The second part is a proper notebook where the learners are firmly invited to write their hypotheses, their comments, etc. Both parts are placed in parallel, so that comments may be associated to some particular event. The activity of describing its own hypotheses and writing comments is based on reflective skills.

Communication will happen through these notebooks: each learner will be able to open the other's notebook, to interrogate him about some note, to answer his questions, etc. We speculate that the ambiguity of the notebook's role, as a tool for communication and for individual reflection, will favour the internalisation of the reflective skills (first, required by communication and then installed as an individual habit).

6. Learning task

The learner's degree of involvement within the interaction will highly influence the actual effects of verbalization and mutual criticisms. We cannot directly tune this degree of involvement because it is essentially an individual parameter, depending on a learner's interest towards the task, towards the co-learner, towards a computer and towards the learning activity in general. Nevertheless, we believe that we have a greater probability of getting adult learners involved with real world problems than with artificial ones. This explains our choice of electoral systems as learning domain.

Two learners may observe electoral simulations in a parallel way. In this case, each learner may have its own task model independently from the other and may carefully avoid any confrontation. A system designed for promoting the negotiation of various viewpoints must include a **task bottle-neck**, i.e. a stage in the process where the concrete activity to perform constrains the learners to agree on a common decision and therefore articulate their views.

In PEOPLE POWER, the task bottle-neck will be to design electoral simulations. Designing a simulation means describing a complete electoral system (parties, candidates, constituencies, rules for expressing preferences, rules for counting votes, rules for allocating seats, ...) and an electoral situation (the preferences distribution in the country). Then the simulation will be run and both learners will observe the results. They will not be directly constrained to exchange their understanding of these results but this confrontation will later be constrained by the task bottle-neck: for justifying her design choice for an experiment, a learner will have to defend her hypotheses and these hypotheses depend on how she interpreted the previous experience(s). Figure 3 illustrates this cycle.

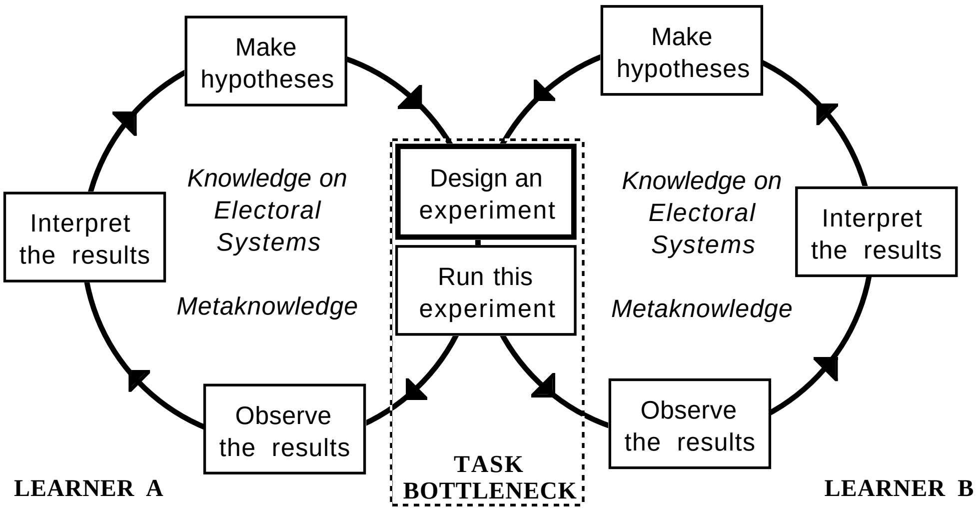

Figure 3: Task bottle-neck in People Power

Since we want to constrain the learners to verbalize metacognitive decisions, the task bottle-neck must require such decisions. Firstly, this means that the learning activity must involve a complex sequence of steps, creating opportunities for strategic decisions, and that is the case in our micro-world.

Secondly, designing each single experiment also requires some metacognitive reasoning since it implies reasoning about the experimental characteristics which are supposed to bring maximal information. In the electoral simulations, the metacognitive knowledge involved includes domain-independent heuristics such as "change one parameter at a time" and context-specific rules-of-thumb, such as "try with a percentage just below 50%".

A third task characteristic involving metacognition is the absence of feedback. The learners will not receive any direct feedback on the correctness of their knowledge. They are responsible for assessing their own knowledge by reflecting on its stability (e.g. knowledge which is confirmed by the last #n experiments is likely to be correct), by communicating with the co-learner or designing further experiments. None of these means will provide them a definitive assessment of the acquired knowledge. This permanently temporary process of assessing one's own knowledge is a critical aspect of everyday metacognition.

7 Design of the co-learner

The educational goals of our system also determine the design of the co-learner component. As is shown clearly in the previous sections, the main focus is given to the learning process rather than to the learning product. More precisely, the expected benefits are related to the interaction during learning, rather than to the quality of the discovered concepts. Furthermore, the quality of this interaction is partially dependent of the **psychological validity** of the co-learner algorithm: if the co-learner's learning method is completely artificial or not understandable for the real student, she will not be able to share this method or discuss it with the co-learner.

These constraints are unfortunately in conflict with the current state of research in machine learning. The original Lancaster project (Gilmore and Self, 1988) aimed to use an improved concept learning technique based on focusing. Such techniques are relatively efficient, but very restricted with respect to the form of interaction they enable: they might simply discuss the fact that an attribute-value pair is present or not in an instance, and that some value is more or less general than another.

Chan and Baskin (1988) met the same problem in the design of "INTEGRATION KID", a collaborative learning system in symbolic integration. The lack of appropriate machine learning techniques led them to abandon these techniques for a pre-defined sequence of skills. The co-learner's evolution is performed through a succession of programs corresponding to different levels of competence. This solution is pragmatic, but brings their system back to the ITS paradigm: a tutor is required to decide when to change the co-learner level, this tutor needing domain knowledge and some student model. Furthermore, the co-learner's behaviour might be only weakly sensitive to the evolution of the learner's behaviour since the interaction between learners has no direct effect on the co-learner competence.

Two elements ease slightly the requirements on the co-learner's psychological validity. Firstly, some minimal difference between learners appears to have a positive effect on collaborative learning. Secondly, 'psychological validity' does not say everything. We may for instance imagine a neural network approach which would reproduce human behaviour but would be unable to explain its results to the other learner. Thirdly, the student’s perception appears to be related more to the mechanisms she attributes to the co-learner rather than to its actual mechanism. Hence, the validity of our choices will not be assessed by collecting and analysing many protocols, but instead through the system's ability to learn efficiently in collaboration with a human learner. We consider below a few directions of our work.

Our main concern in designing the co-learner model is to have a **higher level of granularity** than machine learning systems usually have. Most machine learning programs are indeed based on one or a few monolithic learning methods. Most similarity-based learning methods use induction and discrimination steps, without using for instance explanation techniques for inducing functional rather than perceptive properties. Explanation-based learning methods often neglect the inductive way of guiding generalisation. Most proposed combinations of these techniques still include a succession of a few major steps, e.g. explanation then induction.

Often, the built systems are mainly based on one of these methods, implemented as a relatively simple structure of relatively complex algorithms. Such structure does not leave place for the frequent discussion of learning choices. We see learning more as the use of **complementary inductive, deductive and analogical hints**, used at different moments of the learning process, conjunctively or successively. The global amount of these different kinds of hints gives an inductive-, deductive- or analogical-colour to the global process. In table 1 we propose as example a fictitious and arbitrary decomposition of a simple learning task.

Goal: *What's the formula for calculating the surface of a disk ?*

Step	Knowledge	Formula
Deductive	A surface is the product of 2 dimensions, say dim_1,dim_2	S= dim_1*dim_2
Analogical	For a square: dim_1 = dim_2= width	
	The width of disk is its diameter	S= $diameter^2$
	Test the hypothesis ---> invalidated	
Deductive	The surface of a circle with diameter X is smaller than the surface of square with side X.	S<$diameter^2$
Inductive	In the instances, a larger disk has a larger diameter.	
Inductive	The surface is directly proportional to the diameter	
Deductive	If A is directly proportional to B, A = x * B	S=x*$diameter^2$ & x<1
Inductive	In all the instances, x = 0.785	S=0.785*$diameter^2$
	This formula is now correct, but we have to make it simpler.	
Analogical	The circumference formula also contains a constant ∏.	
Deductive	∏ = 3.14 --> x = ∏/4	S=(∏/4)*$diameter^2$
Deductive	(a/b)*c = (a*c)/b=a*(c/b)	S=∏*($diameter^2$/4)
Deductive	$a^2/b^2 = (a/b)^2$	S=∏*$(diameter/2)^2$
Deductive	Diameter / 2 = Radius	S=∏*$radius^2$

Table 1: A microscopic view of learning.

This simplified instance requires some comments:

- This sequence has been arbitrarily designed, several other sequences could probably lead to the same result. The advantage of a microscopic view is that the system's ability to adopt various learning sequences is inversely proportional to the size of its learning steps.

- Most of the steps in this sequence must be decomposed into smaller steps. For instance, for applying the formula (a/b)*c=(a*c)/b=a*(c/b) to the formula S=(∏/4)*$diameter^2$, it is first necessary to draw an analogy between (a/b)*c and (∏/4)*$diameter^2$.

- This kind of sequence is certainly not defined a-priori by the learner but built progressively during the learning process. This relates this point with the opportunistic/deterministic issue (discussed later on) and the importance of metacognitive knowledge.

In other words, we think learning should be analyzed **at a microscopic level, as a very complex and built "on the fly" sequence of small deductive, inductive or analogical steps**, guided by very important heuristic knowledge. Some steps could be performed by running small and transparent algorithms, especially in induction. The sudden conjunction of different hints could correspond to the unimplemented *insight* process, i.e. the sudden emergence of a bit of information which does not result from a linear reasoning process.

We found some confirmation for the psychological plausibility of this approach in the observations made by Shrager and Klahr (1986) among children learning to use a progammable toy (BIG TRAK). After about 20 minutes, the learners had produced, on average, about 400 key presses which, according to the authors' analysis, corresponded to about 50 learning steps (episodes). Moreover, there was no strict order in the sequence of acquired items, even if some items tended to be acquired earlier, and others later.

The computational feasibility of this approach seems to find support in the partial theory *of plausible reasoning* recently proposed by Collins and Michalski (1989). Their theory includes only four basic transformations: generalization, specialization, similarity and dissimilarity. These relations cover the basic inductive, deductive and analogical operations we listed above. Described in general terms, these relations are associated with nine *certainty parameters* which determine more precisely the actual relation between pieces of knowledge. These are parameters such as the *conditional likelihood* (e.g. how many objects with such characteristics are member of this class), the *frequency* (e.g. how many objects of this class have such characteristics) or the *typicality* (e.g. an ostrich is not a typical bird). Actual implementations of this theory have already been performed, for instance by Baker et al. (1987).

Our microscopic approach has several implications which we consider now.

<u>First implication.</u> In a deterministic approach to learning, the learner designs an experiment according to the information he wants to obtain. This design is partly driven by metaknowledge such as "change one parameter at a time". Hence, our emphasis on metacognition should favour this approach. But, on the other hand, the collaborative setting suggests an **opportunistic** approach because, when the co-learner suggests running some electoral simulations, the real learner will have the freedom to reject the co-learner suggestions and design her own experiment. Consequently, the co-learner may assist in unanticipated experiments. It must be able to learn also from these unanticipated experiments. The term "opportunistic" refers to the situatedness (Suchman, 1987) of learning actions, i.e; the fact that they depend on environmental events as much as on learner plans.

<u>Second implication.</u> AI research has been divided into several disciplines for historic reasons but the human brain did not undergo the same partitioning. The discrimination of learning and problem solving activities has led to several mistakes, such as a dramatic underestimation of the role of the learner's goals and of the context of learning (e.g. the link between the learned concept and the expected performance improvement). The learning situation has been neglected in machine learning as the premiere source of background knowledge. Only during the last few years has this issue become more popular with developments in *explanation-based learning* (DeJong and Mooney, 1986), *purpose-directed analogy* (Kedar-Cabelli, 1986) and *contextual concept learning* (Keller, 1986).

On the other hand, the learning potential of problem solving techniques has also been neglected. If we consider our daily life, we acquire knowledge through many activities which are not considered to be deliberate learning activities. The level of granularity we proposed lies below the discriminations between major mental activities: inducing for learning or inducing for solving a problem both trigger the same basic inductive operation. Globally, the task of the co-learner is a problem solving task: it has to find and explain electoral biases. Some operations performed in problem solving produce durable knowledge changes and then deserve the "learning" label. In other words **learning is viewed here as a side-effect of problem solving**.

Let's consider analogical problem solving. Several authors (reviewed by Hall, 1989) tackle the issue of recording the built analogy, for instance as a new production rule (Anderson and Thompson, 1986). The built analogical link must be considered as a new class, subordinating the source and target concepts, and the knowledge must be reorganised to integrate this new class. If we need for instance to map "Ki: computer directory" with "Kj: table of contents", we need to form the concept "K{ij}: list of subdivisions". Implementing learning as a side-effect implies keeping K{ij} in memory and reorganising Ki and Kj according to their link with K{ij}:

K{ij}: list of subdivisions of a memory
Ki: instance of K{ij} ; subdivisions = files ; memory = disk
Kj: instance of K{ij} ; subdivisions = chapters ; memory = book

<u>Third implication</u>. It is clear that collaborative learning covers a large range of learning methods, including learning by induction, instruction, imitation (apprenticeship), explanation, experimentation, etc. We cannot envisage having a repertory of learning algorithms where the co-learner would select the appropriate one. We think the microscopic approach to learning transcends machine learning terminological distinctions (as we showed for AI topics). The learning performed by a problem solving program will be closer to one or another learning algorithm according to the moment, the data or the response of the real learner.

<u>Fourth implication.</u> So far, we have described a learning model which is **not specific to the collaborative setting**. There is no reason to design a collaborative learning model different from a learning-alone model: students do not change their brain when they learn with somebody. We will not implement some special architecture of collaborative mind. We will rather proceed inversely: the model will have some plausibility if its actual effects in a collaborative setting match some effects observed by psychologists with human pairs.

Let us for instance consider the benefits related to the activity of giving explanations. In her meta-analysis, Webb (1989) found that building an elaborated explanation was related to better achievement, while giving low-level help (an answer, an atomic piece of information) was not. These results are consistent with those collected by Chi et al. (1989) about self-explanations (see section 4.2). Both refer to the same psychological process by which some problem solving mechanism, namely building an explanation, has a side-effect on the explainer's knowledge.

Another example is found in the experimental data presented by Klahr and Dunbar (1988). These experiments are intended to provide data supporting a model of scientific reasoning that they call "dual search space". They compare two experimental conditions. In both conditions, students have to perform experiments with the BIG TRAK device. Their data confirm the natural trend of students to ignore disconfirming evidence and to conserve hypotheses contradicted by the experiments (in more than 50% of the cases). The point which interests us is that these authors observed that, when students are asked to state several hypotheses before the experiment, they are less likely to stick to the disconfirmed hypothesis. In other words, the availability of alternative hypotheses promotes conceptual change. The model they proposed (which is close to a computational form) is based on data collected in learn-alone setting. However it provides one explanation for the conceptual change (Amigues, 1987) mechanisms activated in collaborative settings since an hypothesis proposed by a member of the pair may constitute an alternative hypothesis for the other member.

8. Conclusions.

We have described the current status of the design process of a collaborative learning system. The educational goals, the collaborative style, the learning task and the co-learner algorithm form a logical chain, organised around metacognitive preoccupations. The constraints concerning the co-learner design involve adopting an approach to learning quite distant from current machine learning research ... and we are really worried about the amount of work this represents! However, the fact that this approach raises many interesting issues convinces us that it is worth continuing to investigate it.

Acknowledgments

Thanks to Nicolas Balacheff for his comments on previous drafts of this paper and for allowing us to analyze the protocols he collected, to Ann Nguyen-Xuan for her comments, to Claire O'Malley for her invitation to this workshop and to all those who bought Pierre a beer during the workshop. This work was partly supported by the UK Science and Engineering Research Council grant GR/D/16079.

References

AMIGUES R (1987) Conceptual change and peer interaction. Paper presented at the Second European Conference for Research on Learning and Instruction. Tübingen, West-Germany, September.

ANDERSON J.R. and THOMPSON R. (1986) Use of analogy in a production system architecture. Paper presented at the Illinois workshop on Similarity and Analogy. Champaign-Urbana. Illinois.

BAKER N., BURSTEIN M.H. and COLLINS A.M. (1987) Implementing a Model of Human Plausible Reasoning. Proceedings of the Tenth International Joint Conference on Artificial Intelligence, Vol. 1, pp. 185-188. Morgan Kaufman. Los Altos California.

BALACHEFF N. (1989) Treatment of refutations: aspects of the complexity of a constructivist approach of mathematics learning.in VonGlaserfeld (Ed), Constructivism in Mathematics Education, Reidel.

CHAN T.-W. and BASKIN A.B. (1988) "Studying with the prince", The computer as a learning companion. Proceedings of the ITS-88 Conference, Montreal, June, pp. 194-200.

CHI M.T.H., BASSOK M., LEWIS M.W., REIMANN P. and GLASER R. (1989) Self-Explanations: How Students Study and Use Examples in Learning to Solve Problems. Cognitive Science, 13, pp. 145-182.

COLLINS and BROWN J.S. (1988) The computer as a tool for learning through reflection. in H. Mandl and A. Lesgold (Eds) Learning issues for Intelligent Tutoring Systems. Springer-Verlag. New-York, pp. 1-18.

COLLINS A. and MICHALSKI R. (1989) The logic of Plausible Reasoning: A Core Theory, Cognitive Science, 13, 1-49.

DEJONG G. and MOONEY R. (1986) Explanation-Based Learning: An Alternative View. Machine Learning. Vol.1, nº2, pp. 145-176.

DEWEY J. (1933) How we think. Death & Company. Boston

GILMORE D. and SELF J.A. (1988) The application of machine learning to intelligent tutoring systems. in J.A. SELF (Ed). Artificial Intelligence and Human Learning. Chapman and Hall, London, pp. 179-196.

HALL R.P. (1989) Computational Approaches to Analogical Reasoning: a comparative analysis. Artificial Intelligence, 39, pp. 39-120.

JOINER R. (1989) Mechanisms of Cognitive Change in Peer Interaction: A Critical Review. CITE Report nº60, Institute of Educational Technology, Open University, UK.

KEDAR-CABELLI, S. (1986) Purpose-directed analogy: a summary of current research. in Mitchell T.M, Carbonell J.G and Michalski R.S. Machine Learning: a Guide to Current Research. Kluwer Academic Publishers. Boston.

KELLER, R. (1986) Development of a framework for contextual concept learning. in Mitchell T.M, Carbonell J.G and Michalski R.S. Machine Learning: a Guide to Current Research. Kluwer Academic Publishers. Boston.

KLAHR D. and DUNBAR K. (1988) Dual Space Search During Scientific Reasoning, Cognitive Science, 1-48.

KNIGHTS S. (1985) Reflection and Learning: the importance of the Listener in D. Boud, R. Keogh and D. Walker (Eds) Reflection: turning experience into learning. Kogan-Page, London, pp.85-90.

LOCHHEAD J. (1985) Teaching Analytic Reasoning Skills Through Pair Problem Solving. in J.W. Segal, S.F Chipman and R. Glaser Thinking and learning skills. Volume 1: Relating instruction to research. Laurence Erlbaum, Hillsdale, NJ,pp. 109-131.

MYAKE N. (1986) Constructive Interaction and the Iterative Process of Understanding. Cognitive Science, 10, pp. 151-177.

NGUYEN-XUAN, A. (1989) Conscience, prise de conscience, métacognition. In J.F. Richard (Ed), Le traitement de l'information symbolique (to appear).

O'MALLEY C. (1987) Understanding explanation. Paper presented at the third CeRCLe Workshop, Ullswater, UK.

PALINCSAR A.S. and BROWN A.L. (1984) Reciprocal Teaching of Comprehension-Fostering and Comprehension-Monitoring Activities. Cognition and Instruction, vol.1, nº2, pp. 117-175.

REIMANN P. (1988) Modelling discovery learning processes in a microworld for optics with adaptive production systems. Paper presented at the Second European

Seminar (Summer Univeristy) on Intelligent Tutoring Systems, Le Mans, France, Octobre.

SALOMON G. and GLOBERSON T. (1989) When teams do not function the way they ought to. International journal of Educational research. Vol. 13, nº1, pp.89-100.

SAUNDERS W.M. (1989) Collaborative writing tasks and peer interaction. International journal of Educational research. Vol. 13, nº1,pp.101-112.

SCHOENFELD A.H. (1987)What's all the fuss about metacognition? in A.H. Schoenfeld (Ed) Cognitive Science and mathematics education. Laurence Erlbaum Associates. Hillsdale. NJ., pp. 189-215.

SCHOENFELD A.H. (1989) Ideas in the air: speculations on small group learning, environmental and cultural influences on cognition and epistemology. International journal of Educational research. Vol. 13, nº1, pp.71-88.

SHRAGER J. and KLAHR D. (1988) Instructionless learning about a complex device: the paradigm and observations. Int. Journal of Man-Machine Studies, 25, pp. 153-189.

VYGOTSKY L.S. (written in the 30', but edited by M. Cole, V. John-Steiner, S. Scribner and E. Souberman, in 78), Mind in Society. The Development of Higher Psychological Processes. Harvard University Press. Cambridge, Massachussets.

WEBB N.M (1989) Peer interaction and learning small groups. International journal of Educational research. Vol. 13, nº1,pp.21-39.

WENGER E. (1987) Artificial Intelligence and Tutoring Systems. Computational and Cognitive Approaches to the Communication of Knowledge. Morgan Kaufman, Los Altos, California.