Social and Cognitive Processes in Knowledge Acquisition

Brian R. Gaine
Knowledge Science Institute
University of Calgary
Alberta, Canada T2N 1N4

Abstract

A model of knowledge-acquisition for knowledge-based systems is developed which presents the acquisition activity as playing an essential and continuous role in skilled performance, rather than as a separate and separable activity. The practical implications of this model for systems design are developed, and recommendations made targeted on monitoring the quality of advice from expert systems and achieving closer integration between the application of these systems and the formation of expertise. The model is developed in depth to generate taxonomies of human knowledge processes and use these to analyze the roles of a wide variety of computer-based systems in supporting these processes. The model is used to highlight strengths and weaknesses in the current state of the art in knowledge representation. This paper provides an overall framework for the variety of knowledge acquisition problems, techniques and technologies discussed in the literature.

1 Introduction

One of the most interesting aspects of research on knowledge acquisition is the sheer diversity of topics and approaches that are reported: from cognitive models of human behavior to computer tools for knowledge elicitation; from interviewing techniques to text analysis; from practical experience of expert system development to laboratory studies of the consistency of expertise; from narrow domains of particular systems to broad requirements for the integration of numeric computing, simulation, databases, communications and hypermedia. However, this diversity is also confusing. What is the underlying cohesive theme of ‘knowledge acquisition.’ Does the diversity reflect the sheer anarchy of human knowledge processes as Feyerabend (1975) suggests, or is there some reasonable infrastructure in common to the different aspects of knowledge acquisition? Is not the whole of computing in the support of knowledge and, if so, what is special about knowledge-based systems and knowledge acquisition?

One perspective that enables us to address theses issues is to take the cognitive entity to be considered in studying knowledge acquisition as human society rather than an individual (Gaines 1987a). Society acts as a distributed anticipatory system modeling and controlling the world to enhance its probability of continued existence. Knowledge transfer processes within society are manifestations of coordinating communications between the components of the distributed system. A taxonomy of such processes may be used to classify various approaches to the computer-based support of knowledge acquisition. Expertise formation is a manifestation of social processes encouraging functional differentiation in the, otherwise fairly uniform, distributed anticipatory system (Gaines 1987b). Our culture provides a wide variety of positive feedback mechanisms, corresponding to the taxonomy of knowledge transfer processes, that: support individuals in the development of excellence; discourage uniformity in specialization; and encourage the sharing of specialist knowledge.

Techniques being developed for knowledge-based systems and knowledge acquisition may be seen as supporting the variety of knowledge transfer processes in society. The use of computing technology to automate these techniques follows the developmental trends in information technology (Gaines & Shaw 1986): from manual techniques in the 1970s; through a first generation of computer-based tools in the early 1980s; to a second generation of integrated knowledge support systems in the late 1980s (Gaines 1988a). The embedding of knowledge acquisition technologies within the general infrastructure of information technology raises many integration issues from person-computer interaction, through the roles of experts, clients and knowledge engineers, to the relation of other computing technologies to the shell, support environment and knowledge acquisition tools. The knowledge structures underlying these issues may be used to generate complementary design requirements at each branch in the taxonomy, one relating to integration issues and the other to separation issues (Gaines 1987c). These issues may also be analyzed in the wider framework of the variety of human knowledge processes, the underlying techniques involved, and pre- and post-computer supporting technologies (Gaines 1988a).

This paper focuses on the grounding of knowledge acquisition techniques and technology in human social and cognitive processes. It attempts to provide an integrative framework for knowledge acquisition for knowledge-based systems that has deep foundations but is also clearly applicable to the technology and systems development.

2 The Embedding of Knowledge Acquisition in Expert Activity

The distinction between knowledge acquisition and performance systems is an invidious one. It can be taken to imply that expertise is a static collection of skills and knowledge, capable of being elicited and transferred to an expert system that will then be able to reproduce the expertise. This static model is an extremely dubious one in systems design, leading to systems that date rapidly, interfere with the growth of expertise, have poor reliability and high maintenance costs. The following model explains how these conditions arise, and leads to approaches for avoiding them.

Hawkins (1983) has abstracted from industrial experience in developing mineral exploration expert systems and proposed a model of human expertise relevant to expert systems. Figure 1 presents a summary of the essential features of this model:

- The expert first elicits data about the problem from the client;
- He or she develops a minimal model that accounts for the data provided;
- He or she generates advice based on the model and feeds this back to the client;
- The client may accept the advice, or query it and, possibly, the model;
- he queries lead to further data elicitation, and repeat of the elicitation/modeling/advice/query cycle.

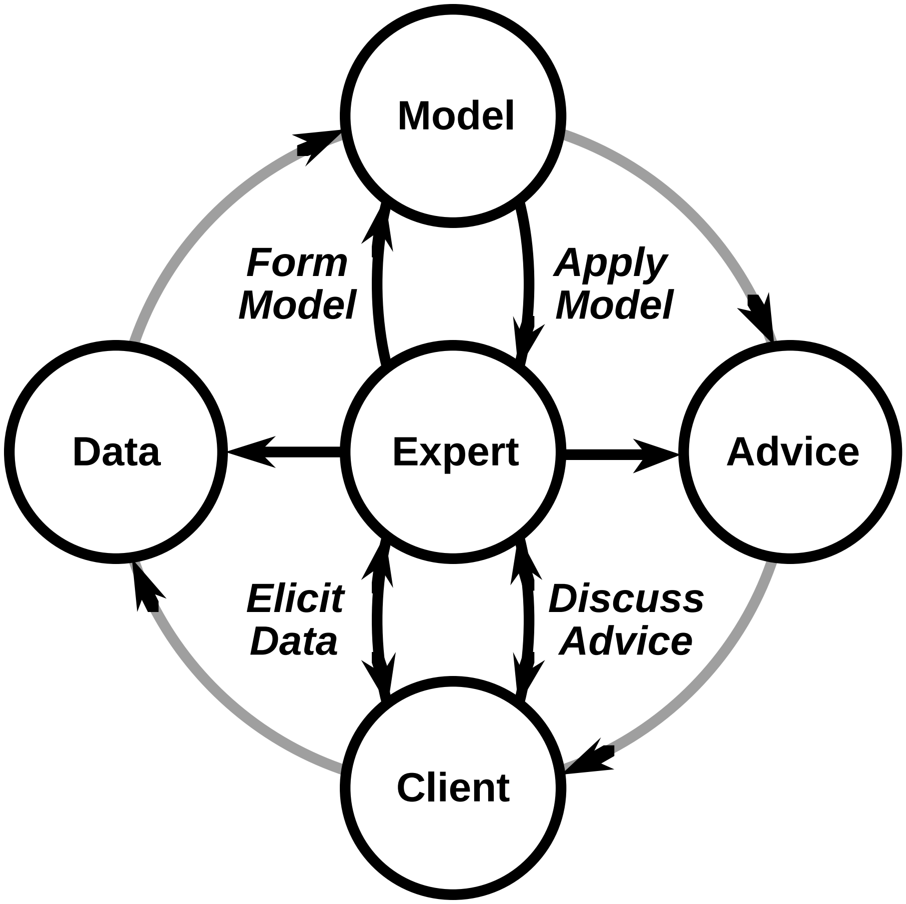

Figure 1 The negotiation cycle in expert-client interaction

Thus, in Hawkins' model, the client plays an active role in further developing the model by providing more data until he or she is satisfied with the model and consequent advice. Expert advice giving and taking is part of a cycle of negotiation around a process of model formation.

The particular importance of this model of expert performance is the emphasis it places on the embedded knowledge acquisition process. The expert is managing a data collection, modeling and decision cycle within a consultation. As Figure 1 shows, this implies that the expert is able to enhance his or her collection of case histories, models and advisory strategies through the consultation. Skilled performance is not the static utilization of expert knowledge and skills in Hawkins' model—it is instead a microcosm, and an essential part of, the wider processes of knowledge acquisition which are intrinsic to the role of an expert.

3 Extending Acquisition through the Professional Community

Access to the learning experiences of attempting to solve the problems of a client community is one major source of knowledge in the development of the expertise of an individual expert. The client community provides access to practical experience, a variety of problems, particularly novel ones that go beyond existing expertise, and it also manages the growth of expertise through systems of reward, criticism and access to resources. However, learning from experience is a slow and error-prone process, and socially significant areas of expertise become associated with professional communities that attempt to expedite learning and reduce errors through the sharing of experience. Figure 2 extends the model of Figure 1 to include these other processes, such as apprenticeship, instruction, training, education, workshops, conferences, journals and books. Professional communities usually also play a major role in refining and directing the client community's reward, criticism and resource allocation systems.

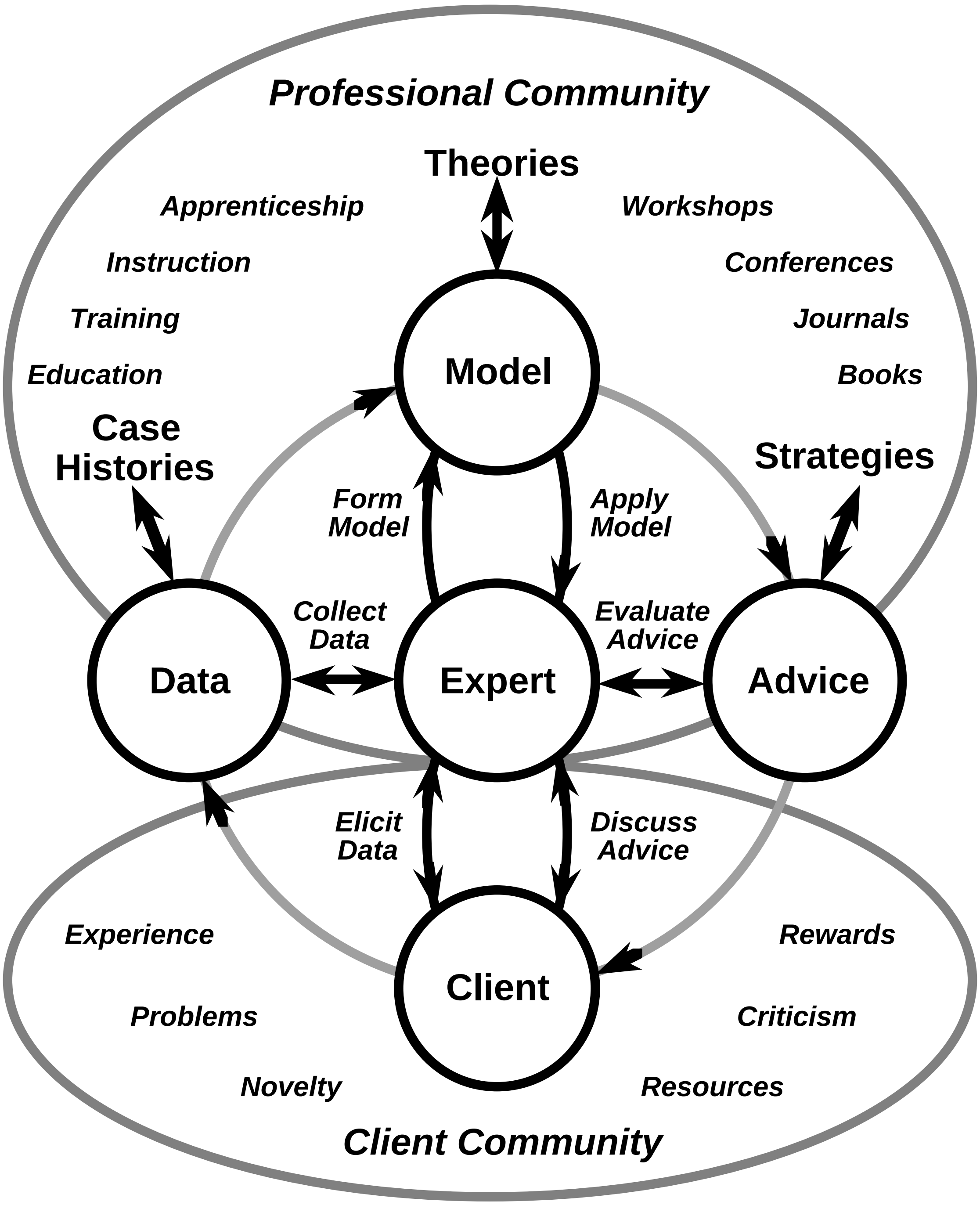

Figure 2 Extended knowledge acquisition process in expert-community interaction

The role of the professional community can be seen as an indirect enhancement of the basic processes of knowledge acquisition in the negotiation cycle of expert performance. The individual can acquire the data of others through access to reported case histories. He or she can enhance his or her repertoire of models through the published theories developed by others to account for their data and the data of many others. He or she can enhance his or her repertoire of advice through the published advisory strategies developed by others and evaluated by them, or yet others, against past case histories, theoretical foundations, and experience in use.

This model of individual expertise as intrinsically a knowledge acquisition process at multiple levels intrinsically embedded in, part of, and supported by essential social processes, is consistent with Stich and Nesbitt's (1984) view of the social role of experts as managers of the *reflective equilibrium* of the inductive process of knowledge generation in society. It is also consistent with the emphasis in professional development books concerned with good practice on the importance of balancing experts' authoritarian roles as knowledge repositories and appliers with their entrepreneurial roles as knowledge acquirers (Schön 1983).

Sociologists have noted the very strong positive feedback processes in the dynamics of knowledge acquisition in the scientific community (Hagstrom 1965). Merton (1973) coined the term the "Matthew effect" for those features of the reward system in research that were biased towards allocating greater credit for the same discovery to those with an already established reputation. The qualitative effect of such positive feedback processes is to amplify random differences within a population, creating strong distinctions which have no other basis than the feedback process itself. That is, even if all scientists were created equal in their capabilities, some would become experts relative to others because random differences in early performance generated credit that affected resource allocation and gave them greater opportunities to acquire knowledge. This strength of this effect can be demonstrated by studies of the behavior of 'experts' simulated by learning automata in a competitive learning environment with access to resources related to demonstrated 'expertise' (Gaines 1988b).

In many areas of expertise where expert systems are expected to play a major role these positive feedback processes are exceptionally strong. In medicine, for example, the key learning resource is access to medical problems, and the 'owner' of such a problem has a strong personal interest in only allowing someone of very good reputation to handle it. The system, including considerations of legal liability, strongly funnels problems to those who are regarded as experts. It is, however, precisely these problems from which new knowledge is generated. Hence direct experience of case history data is made available primarily to those with already established reputations for having capabilities to deal with it. Similar considerations apply to the award of scholarships, invitations to scientific congresses, and so on (Blume 1974). They also apply not only to individuals but also to social units, such as a company subject to government procurement procedures that are heavily biassed to contractors with prior experience and with whom the government agency has prior experience.

4 Implications of Social Acquisition Models for System Design

The model of the expert's relation to those seeking advice as one of trading it for access to problem-solving, and hence learning, opportunities is also consistent with the actual behavior patterns of experts in relation to expert systems. Divesting the unprofitable trading to focus on the profitable is a sensible strategy, that is, passing to an expert system those standard problems where little learning opportunity is expected. However, the uncertainties about the classification of any problem situation suggest there may be dangers in this approach which could be avoided by adequate design and application of knowledge-based systems. The advice giving system should contain rules which indicate problem-solving situations that are beyond the boundaries of its capabilities and refer these to a human expert (in the same way that human experts refer difficult problems to others). It should log actual case histories of advice giving for later review by human experts so that these rules can be validated. It should make provision for feedback as to the results of advice so that learning opportunities are not lost.

The continuity between the modeling role adopted by the expert in Hawkin's model of expert-client negotiation, and the expert's wider modeling role in managing the inductive inference process, suggests that the performance and knowledge acquisition phases of expert systems should not be separated as they are currently. It would be better to think in terms of *knowledge support systems* that integrate knowledge acquisition and advisory roles. In particular, the model of system-user relations as mutual negotiation rather than authoritarian advice giving is a good model for knowledge-based system.

The emphasis on the natural unit representing expertise being a social one rather than an individual emphasizes the need for knowledge-based system development to be based on groups of experts rather than individuals. Figure 2 illustrates the diversity of sources, models and strategies likely in such a group and emphasizes the unlikelihood of there being a complete *consensus* in the knowledge used by the experts. The ongoing acquisition process makes consensus not only unlikely but also undesirable since competing, possibly mutually contradictory, interpretations, models and strategies are an essential feature of active knowledge development. It is important to develop knowledge-based systems that can encompass multiple sources of expertise that may be mutually conflicting, and make effective use of the conflict, for example, to generate 'dissenting' opinions that indicate the boundaries of the expertise (Boose & Bradshaw 1987). Knowledge acquisition systems must be able to elicit, make explicit, and make effective use of different conceptual systems and terminologies (Shaw & Gaines 1988).

All these considerations are also relevant to questions about the accountability of expert, expert system, and client for the advice. However, that is a side-effect of the social pressures on the inductive process—accountability, and the resultant rewards and punishments, are cultural mechanisms to encourage learning and the appropriate use of expertise. Well designed expert systems will make provision for the clear presentation of the quality of advice, reference to human experts, audit trails, logging, and case history review, because these are necessary to an effective technology—the possible legal consequences of not doing this are just an indication of social expectations for an effective technology in the domain of knowledge acquisition and transfer.

There are implications for the detailed technical design of knowledge-based systems. As an example consider the form of rule often induced from case data which states that if the value of an attribute is unknown it is reasonable to fill in it with a default value based on the value of other attributes; for example, if the induction algorithm relates the value of attribute C to that of attribute A:

If (A Equal B) And (C Equal Unknown) Then (Set C Equal D)

This rule uses known data and induced relations to make suppositions. It is also reasonable, but unusual, to add a meta-rule based on the same relation:

If (A Equal B) And (C Not-Equal D) Then (Set Anomaly Equal True)

This meta-rule uses known data and induced relations to highlight anomalies (unexpected data). The new attribute 'Anomaly' indicates that the value of C is inconsistent with previous data used in inductive inference. It can be used to trigger query or alarm reports, and the normal explanation mechanism based on back-tracking will trace the anomaly to its source (Rappaport & Gaines 1988). It can be used to indicate that the problem under consideration is unusual and should be referred to a human expert or, at least, be recorded and flagged of interest.

In this example, the conflict between the rule and the data is exactly the reflective equilibrium managed by experts noted above. The meta-attributes generated through anomaly-detection rules may be used as a basis for machine learning, as Lenat (Davis & Lenat 1982) has done in AM (terming anomalies 'interesting'). However, they are also a basis for effective integration between expert and expert system, with the human expert continuing to have the role of managing the inductive process.

5 Socio-Cognitive Foundations of Knowledge Processes

The extended knowledge acquisition process model of Figure 2 is attractive in capturing much of what we know about expertise and its role and formation in society and the individual. It suggests a number of significant design considerations in the development of knowledge-based systems as discussed in the last Section. However, it is not detailed enough to act as a design framework in its own right. This Section develops the model in greater deal and then uses it to further discuss the role and development of knowledge-based systems.

Figure 3 shows the basic cognitive unit normally considered—the individual as an anticipatory system, experiencing the world, perceiving it, acting upon it, modeling it to predict and explain it, interacting with it to experiment with and control it. The cognitive structures developed are valid in so far as they correspond to, or apply to, the world.

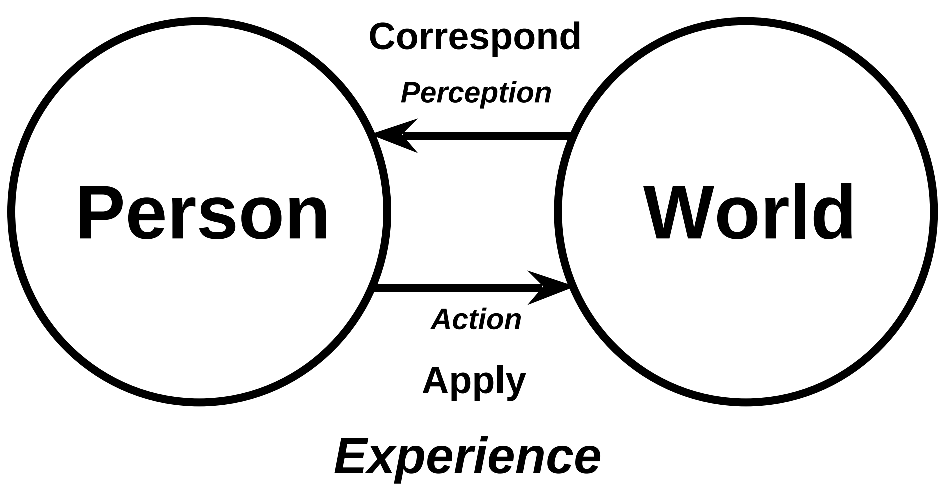

Figure 3 Person as an anticipatory system interacting with the world

I have previously used a hierarchical structure based on Klir's (1985) modeling theory to analyze the activities of the individual in coming to terms with the world (Gaines 1987a). Figure 4 shows a variant of this hierarchy emphasizing the cognitive processes involved in modeling the world and their definitions in abstract, systemic terms:

- To *recognize* at the lowest level is the capability to notice recurrence of the ‘same’ events in the world when they recur. This is already a significant cognitive act because ‘same’ is subject to personal definition, and the concept that events recur is a strong presupposition. Recognition is fundamental to any modeling system but is in itself a weak operation since it is dependent on the recurrence to make use of the data.
- To *recall* at the next level is the capability to regenerate the distinction used in recognition internally so that it is itself an ‘event’ that may be processed. This facility to recreate events in the ‘imagination’ is fundamental to the existence of the cognitive process, detaching human knowledge processing from the immediacy of experience.
- To *model* at the next level is the capability to derive the distinction used in recognition and recall from other distinctions that may themselves not relate directly to experience. This facility to ‘explain’ events in terms of ‘deeper’ distinctions that only indirectly relate to experience is again fundamental to the efficiency of the cognitive process, allowing novel distinctions to be developed that efficiently encode wide ranges of otherwise unrelated experience.
- To *reconstrue* at the next level is the capability to derive one distinction from multiple models. This facility to move between modeling systems is fundamental to the adaptability of the cognitive process, and the human species, allowing a wide repertoire of anticipatory sub-systems to be developed to cope with the variety of the world.
- To *abstract* at the next level is the capability to detach distinctions from their sources and make the relations between models themselves subject to study. This facility to study the world of modeling as if it were a world of experience is fundamental to the externalization and growth of human knowledge as a cumulative by-product of the anticipatory process. It makes the expertise of the species largely independent of that of existing individuals.
- To *originate* at the next level is the capability to treat the distinction making process itself as a human activity, subject to choice and change, and to generate distinctions in themselves. This freedom raises many questions as to the nature of ‘reality’, of the ‘wisdom’ of certain distinctions, and of the relationship of the distinction-making, cognitive and knowledge processes to the nature of the individual and the survival of the species.

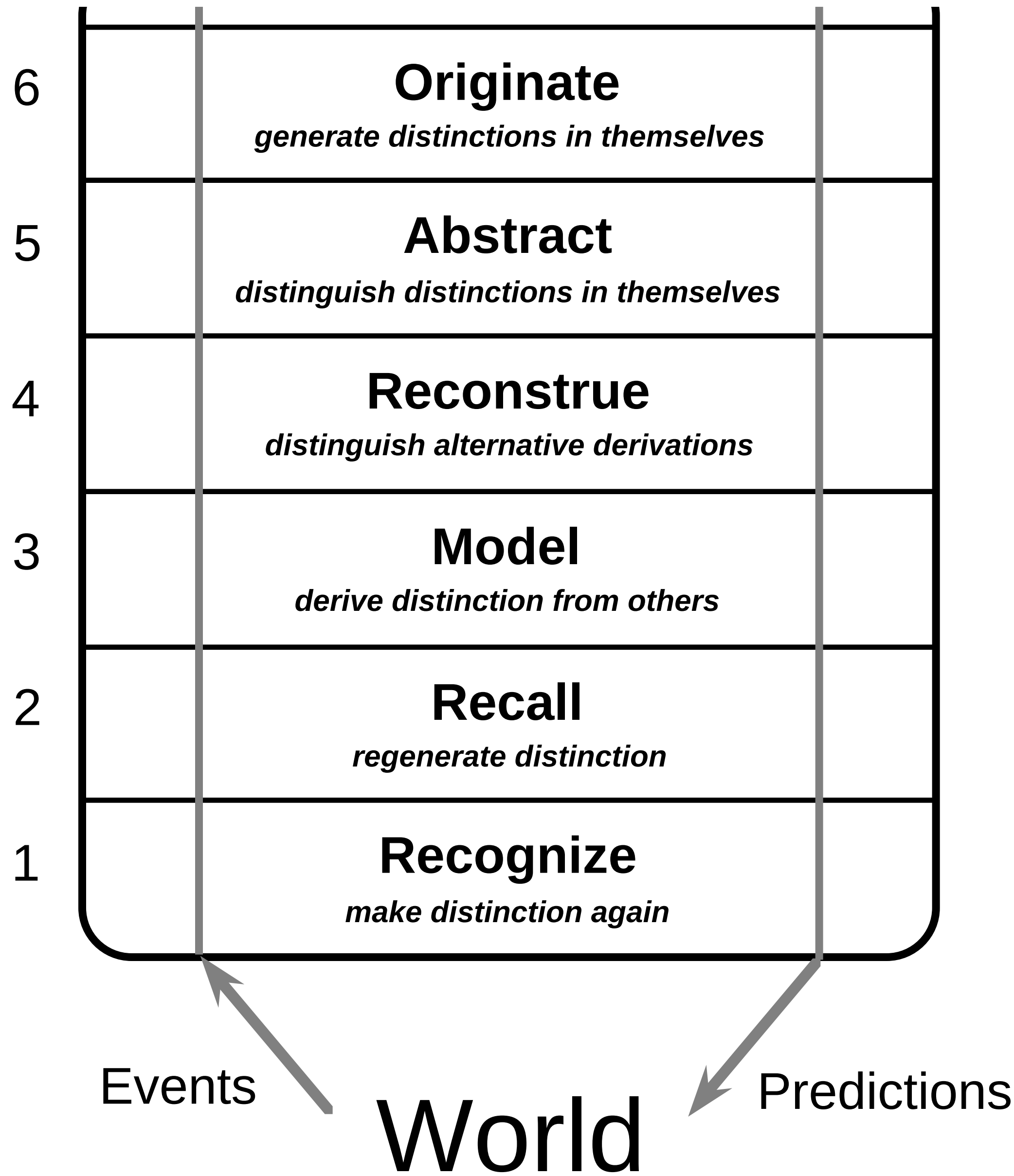

Figure 4 Anticipation as a hierarchy of distinction-making processes

The anticipatory process encompasses interaction as much as it does modeling—we can anticipate the world by predicting where it is going or by controlling it to go a way with which we are equipped to deal. The combined hierarchies along the dimensions of increasing knowledge and increasing interaction give an operational model of power in the normal sense of the word, as the potential to change the course of events. Figure 5 shows the hierarchy of Figure 4 paralleled to encompass interaction with the world and used as a model of power.

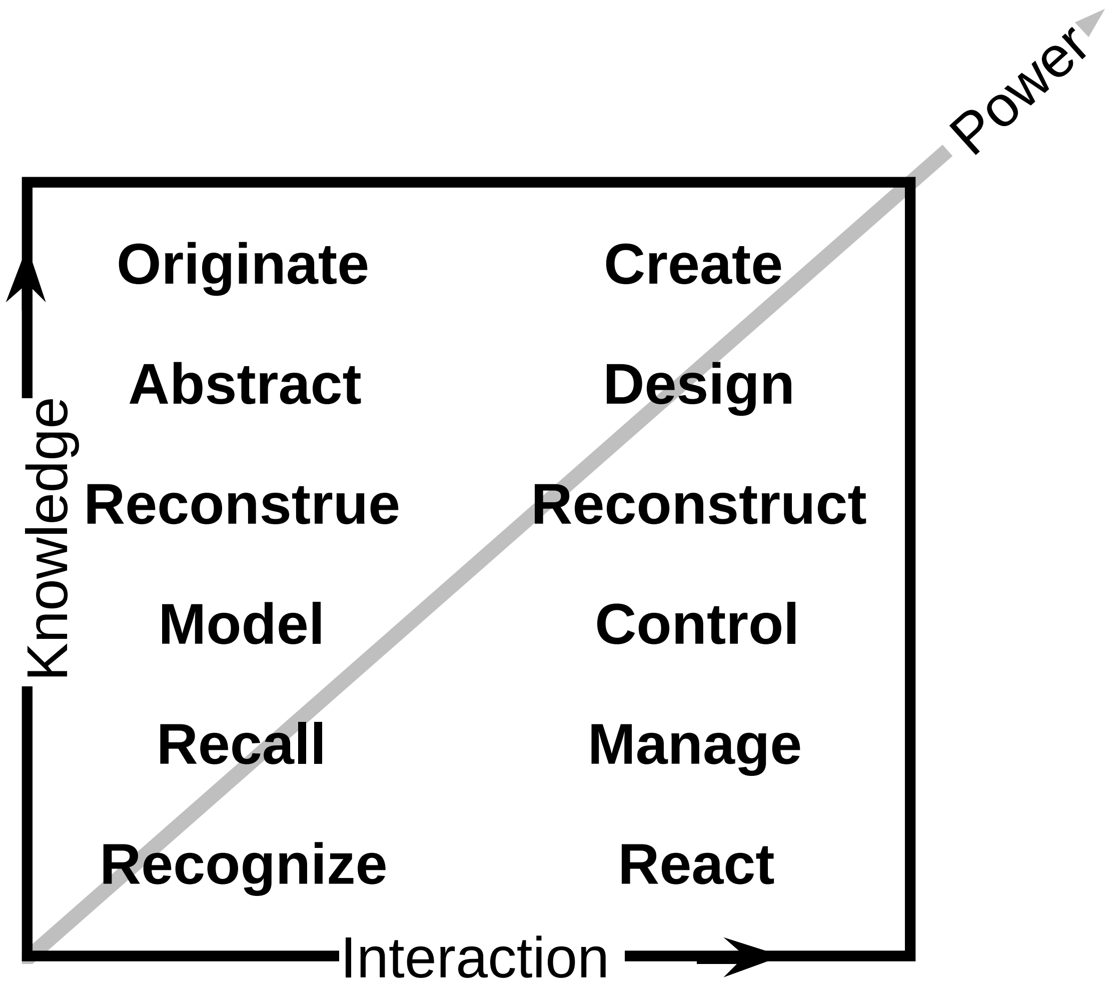

Figure 5 Interactive anticipation as a hierarchy of distinction-making processes

In terms of the interaction hierarchy:

- To *react* at the lowest level is the capability to respond to the same event again in the same way.
- To *manage* at the next level is the capability to regenerate reactions and use them outside their original context. This capability of managing situations has been noted by Hayes-Roth (1984) to be particularly suited to knowledge-based system precisely because it does not involve the use of deep models.
- To *control* at the next level is the capability to derive reactions from models.
- To *reconstruct* at the next level is the capability to change the models on which reactions are based.
- To *design* at the next level is the capability to derive the models from abstract distinctions.
- To *create* at the next level is the capability to generate new models and reactions.

6 Physical, Mental and Knowledge Worlds

One of the most important distinctions made in modeling experience of the world is to separate the world of other people from the physical world. Other people are recognized by the individual as analogies of the self, themselves anticipating the world, modeling and controlling it in a way unlike causal physical systems. In addition other people act as role models, knowledge sources, and creators of new knowledge, in a way that makes the dimensions of interaction with the social world very different from those of the physical world. Figure 6 recognizes this separation and shows the person having communication with a *mental* world in which perception involves the comprehension of communication, and action involves the capability to generate comprehensible explanations. I have used the term ‘mental’ very deliberately instead of ‘social’ or ‘cultural’ world because it emphasizes the lack of separation between the mental world of the individual and the equivalent phenomenon in society. The distributed anticipatory system which is the species is best regarded as a unitary knowledge processing entity, with individual knowledge processes differentiated and independent from the whole to a very small extent compared with their integration with, and dependence upon, that whole.

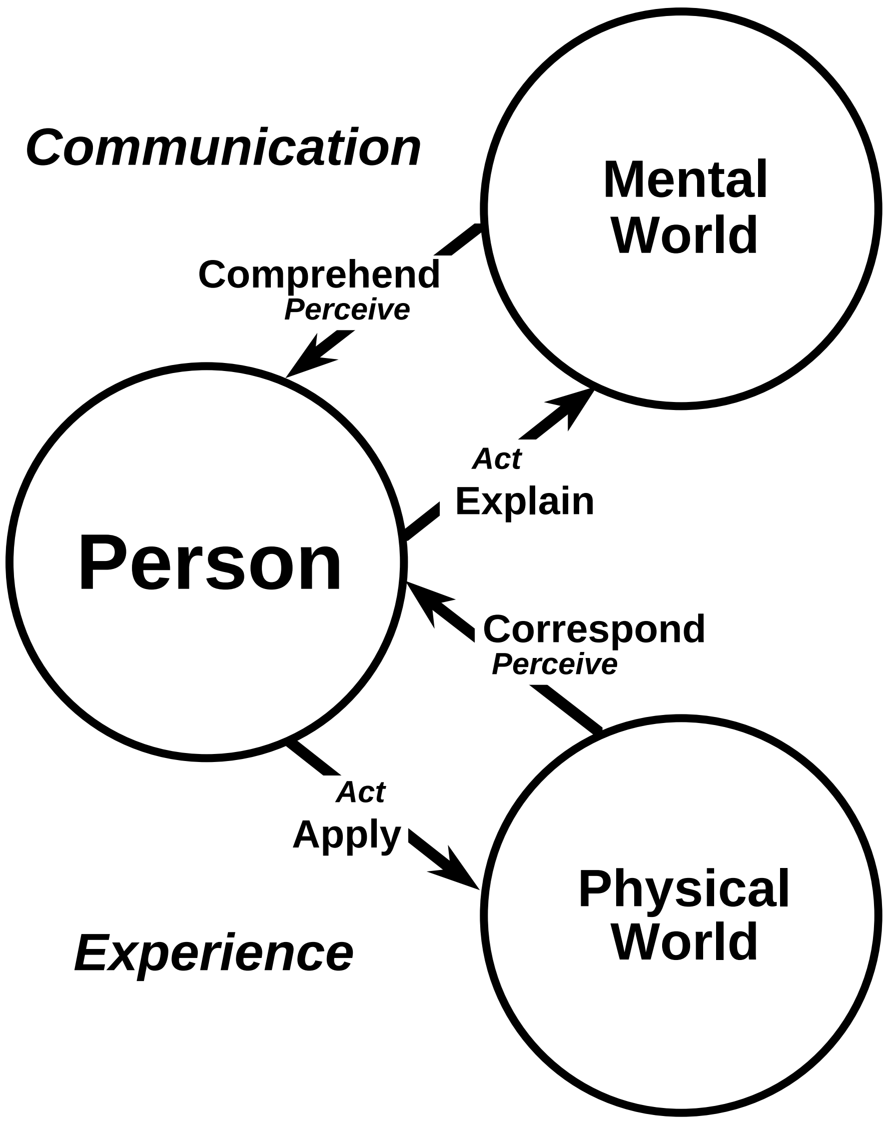

Figure 6 Person as an anticipatory system communicating with the mental world

Figure 6 also emphasizes the perceptions of and actions upon the mental world are evaluated in a different way from those relating to the physical world. It is possible to collapse the mental into the physical and treat the mental processes of people as if they were the physical processes of objects, attempting to predict and control them. Habermas (1984) terms this ‘strategic’ social action, involving including the decisions and behaviors of other people in a means-end calculation. Social actors in the strategic mode relate to other persons as objectifiable means or obstacles to the attainment of their aims. However, the mental world is one of human choice within a framework of established conventions and hence different modes of interaction with it are possible. It is, for example, important to communicate to others what family of conventions one intends to adhere to. Habermas terms this ‘dramaturgical’ social action involving the projection of a public image. It is also important to operate together as a social unit within a common framework of conventions. Habermas terms this ‘normative’ social action in which the parties involved fulfil reciprocal expectations by conforming their behavior to shared norms and values.

Habermas’ fourth category of social action, *communicative action*, is particularly interesting because it involves the persons involved attempting to reach an explicit agreement on their situation for the sake of cooperation. This introduces the concept of mutual understanding and hence of mutual knowledge suggesting a further distinction that separates the shareable distinctions involved in communicative action from the less shareable distinctions involved in individual activities and social interaction without mutual understanding. Figure 7 extends Figure 6 to make this further distinction, showing the person interacting with a ‘knowledge world’ also, in which perception involves the derivation of knowledge from other knowledge, and action involves the capability to create new knowledge.

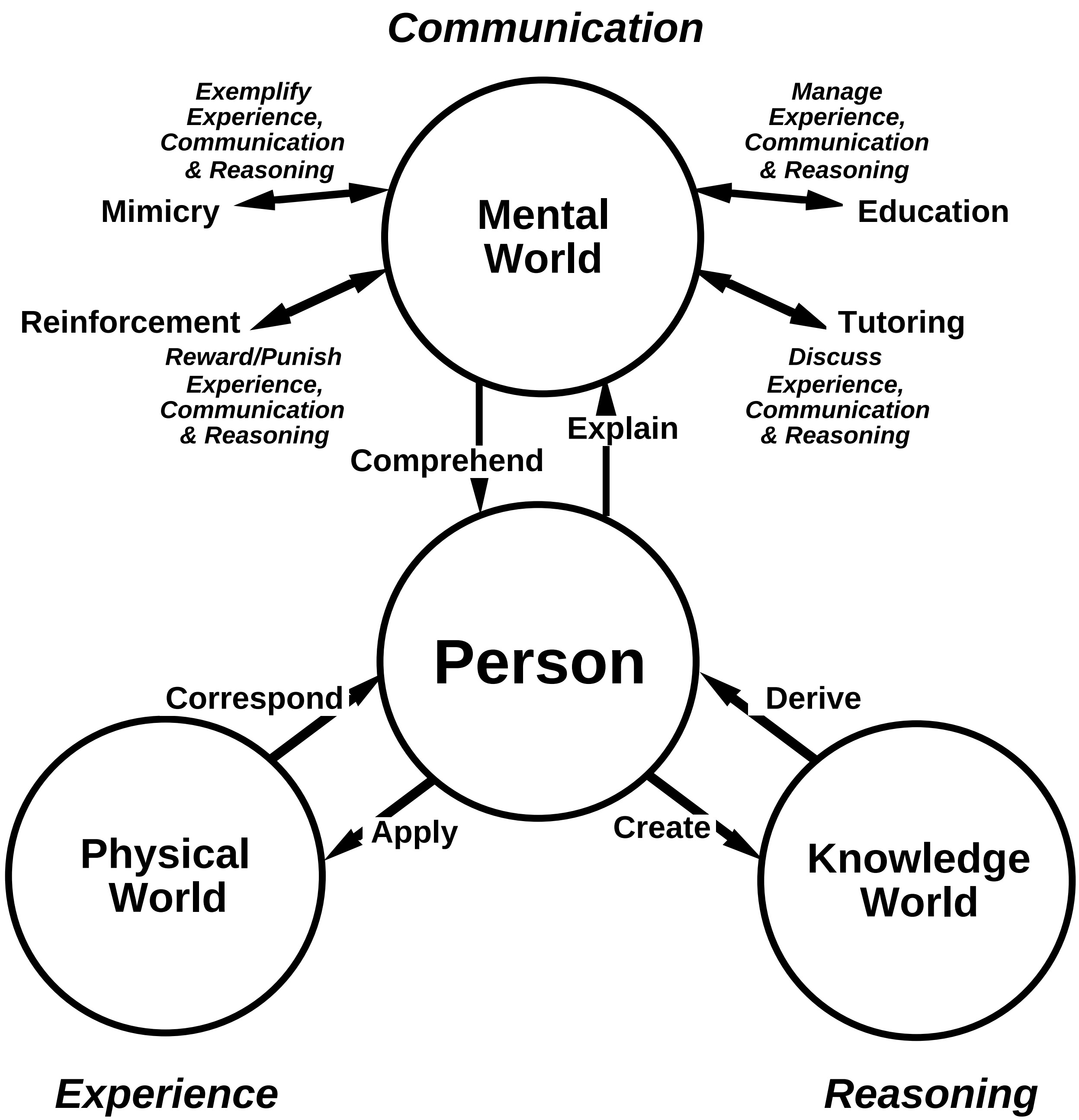

Figure 7 Person as an anticipatory system reasoning in the knowledge world

This knowledge world has many of the properties of Popper's (1968) world 3 of *objective knowledge*. In particular, it is a product of the human mental world, but distinguished just to the extent that it stands independent of the particular mental worlds from which it was created. There is an important conceptual link between Popper's criteria for objective knowledge and the existence of a social world valuing communicative action. The primary criterion is one of *fallibilism*—there must be tests that might show the knowledge to be false. However, this criterion which establishes the possibility of a critical community, also establishes the need for knowledge to be *understandable* so that it can be tested by the community (Notturno 1985).

Pask (1980) has developed a similar framework to Habermas as a foundation for computer-based learning. His *conversation theory* singles out as the critical event in discourse the occasions on which two participants agree that they have effectively exchanged a concept and validated its mutual similarity. Such *agreement over an understanding* punctuates the discourse into *strict conversations* whose structure and processes can be analyzed in depth as coherent units. Pask recognizes the diversity of validity tests that can be applied to understanding such as the ability to: *produce* knowledge transferred; to *reproduce* knowledge transferred from an alternative derivation; to *apply* knowledge transferred to problem solving; and, recursively, to *teach back* knowledge transferred. These criteria correspond to activities at different levels in Figures 4 and 5, instantiating the means open to people to undertake Habermas' communicative action through attempting to reach Pask's agreement over an understanding using Popper 's objective knowledge.

Figure 7 also shows some of the ways in which activities at different levels of the hierarchies of Figures 4 and 5 are supported through the social world. The reflexive capability to react to the physical world is paralleled by a corresponding capability to *mimic* in the social world. This is an important knowledge transfer mechanism that is supported by social action in exemplifying experience, communication and reasoning. The reinforcement processes in the physical world supporting the development of second level capabilities to recall and manage based on evaluated experience are extended by social reinforcement processes. The derivation of relations between distinctions at the third level of modeling and control is supported by explicit tutoring discussing experience, communication and reasoning.

It is important to note also that all these processes may be applied recursively to enhance their own operations. We can mimic and reinforce those who are effective as exemplifiers, reinforcers and tutors. In particular, as indicated in Figure 7, we can manage the process of knowledge transfer through the educational system, although our knowledge of the acquisition process is such that we cannot yet go beyond this to control, reconstruction, design and creation.

7 The Role of Knowledge-Based Systems

Figures 5 and 7 have two important inter-relations. The development of the distinctions between the physical, mental and knowledge worlds shown in Figures 3, 6 and 7, may be seen as a necessary component of the ontogenesis of the hierarchies of Figures 4 and 5. From an individual perspective, the person alone cannot develop this hierarchy without the support of the social infrastructure of other people, and hence the shared mental world is an important construct to be identified as separate from the physical world. We can, and do, of course participate within the socio-cultural world without explicitly modeling it and its effect on us, but the full freedom of the upper levels of the hierarchy involves the recursive application and extension of our models to include the routes by which we arrived there. The separation of the knowledge world is a consequence of our modeling the cultural mental world as separated from our individual mental world and residing in the minds of other individuals, that is, of imposing our models of our selves upon the knowledge processes of society. It corresponds to the detachment of distinctions from experience at levels 3 and above in the hierarchies of Figures 4 and 5.

From a distributed anticipatory system perspective, the separation of the individual and cultural mental worlds is an artefact of the communications interfaces and the limits of individual rationality. Conventional views of the individual mental world and the knowledge world result from our individual-centered perspective on the survival and evolutionary processes of the species. If we see the species as a the unit then the cultural mental world is the manifestation of increasing power and the knowledge world is a product supporting that power.

This is important in understanding and evaluating the role of computer technology in supporting human knowledge processes. Computer technology has largely supported the more abstract human knowledge processes in the past such as those involving modeling, mathematics and calculation. Expert systems may be seen as an attempt to support lower level processes where models have not been available and the tasks have had to be 'managed' by people. Robotics may be seen as as attempt to support even lower level processes of interaction with the world.

This move towards technological support of the lower levels when the growth of knowledge emphasizes the upper levels may appear surprising. However, it is not so. First, the primary pattern of development in our civilization has been to the release of human labor from the physical necessities of survival to the freedom and creativity of experiental, social and intellectual activities. The mechanization of agriculture and the automation of factories are major developments that are now being paralleled by the computer support of low-level mental and physical functions. Second, the development of the upper levels of the anticipatory hierarchy through inductive processes requires an increasingly massive flow of experience through the lower levels. The *grounding* of our knowledge in experience requires a wealth and diversity of experience, and a level of flow of communications, that the biomass alone can no longer encompass. Expert systems are one further step towards the enhancement of human experience through instrumentation and interaction technologies.

In these terms, knowledge-based systems in general may be distinguished from expert systems in particular through their integration of knowledge processes at all levels of the hierarchy. The term 'expert systems' arose because of the emphasis on a new paradigm of modeling the human operator managing those decision and control situations which are not amenable to modeling. That is, expert systems developments have focused precisely on those situations where our knowledge is only at level 2 in the hierarchies of Figures 4 and 5. Thus, it is not surprising that problems of knowledge transfer from human experts arise in the development of expert systems—these are the same problems that prevent the expertise involved rising from recognition and recall to modeling, from reaction and management to control.

In terms of these considerations, we should beware of two extremes. First, of assuming that all human expertise is of this nature—it encompasses the complete hierarchy of Figures 4 and 5. Second, of treating expert systems developments at level 2 as of lesser significance than the upper, more 'scientific', levels of the hierarchy. To the contrary, these developments may be viewed as part of a movement towards the major advances in the flow of information between levels 2 and 3 of the hierarchy, prepatory to similar advances between levels 1 and 2. Moving along the 'power' dimension of Figure 5 involves a very high level of activity at the lower levels if that at the upper levels is to be grounded in reality and applicable to the physical world.

8 The Evaluation of Knowledge-Based Systems

One immediate application of the concepts developed in Figure 7 is to the evaluation of knowledge-based systems, an important topic which is being increasingly addressed by empirical studies (Shaw and Woodward 1987) and the development of conceptual frameworks (Benbasat

& Dhaliwal1988). In expert system developments, one naturally thinks in terms of validation against the expert, either against his or her knowledge explicitly expressed, or against his or her behavior. However, these ‘subjective’ dimensions of validation omit more conventional dimensions, such as: the evaluation of the overall performance of the system in applications; the correspondence of the knowledge in the system to empirical data; the derivation of the knowledge from established scientific knowledge; and the comprehensibity of the knowledge to the professional community as opposed to the individual experts.

Figure 8 shows the person in Figure 7 replaced by a knowledge-based system, together with links to the validation criteria used in expert systems and positive science, and with the ‘worlds’ filled in more formally in terms of relevant concepts and definitions. This diagram gives a framework allowing all six of the dimensions of validation mentioned above to be clearly defined, together with three others that may be less obvious.

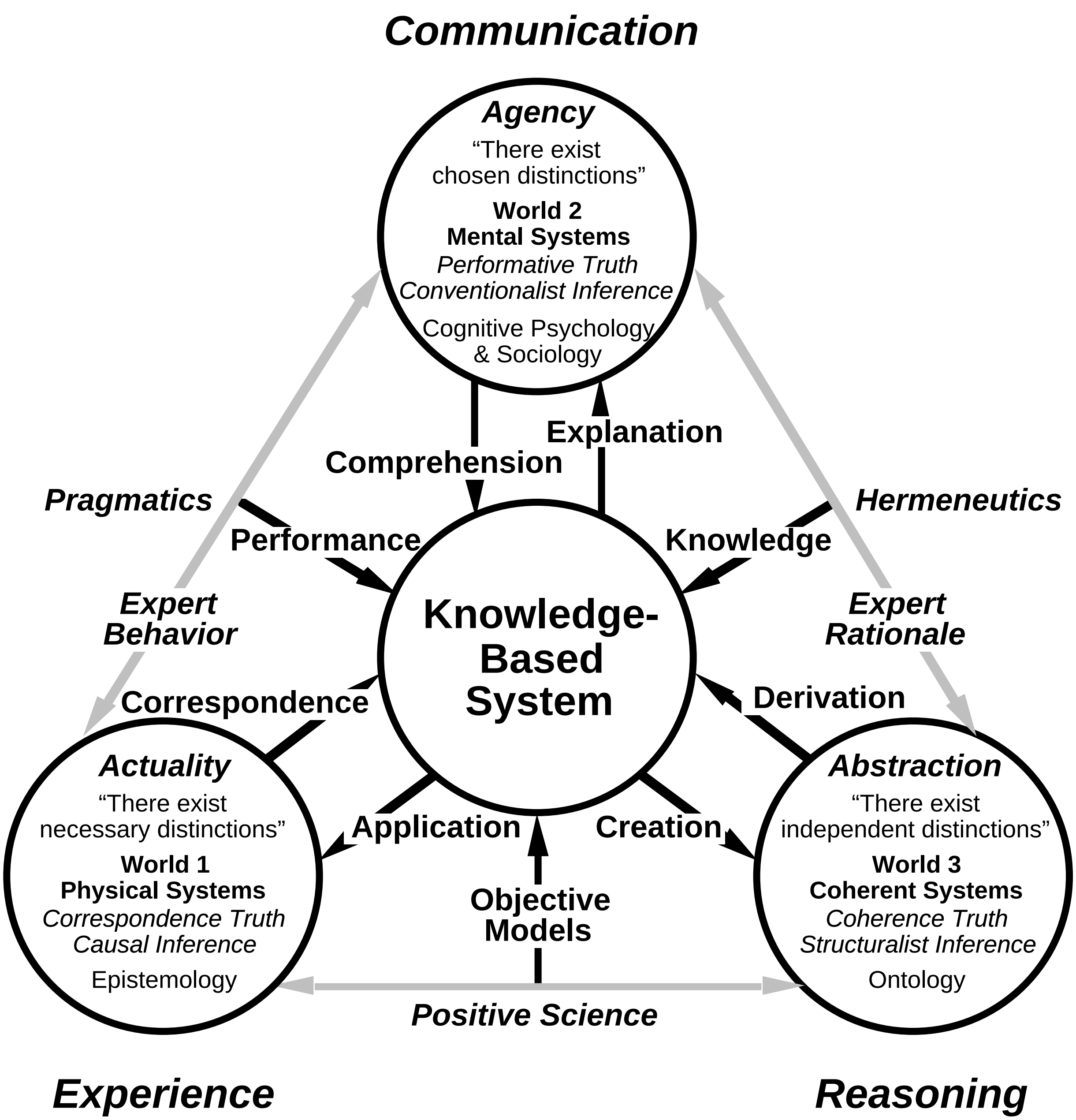

Figure 8 Dimensions of evaluation of a knowledge-based system

The interactions of the system with some actual world in which it is intended to perform decision or control tasks introduce two dimensions:

- *Correspondence*—to what extent does the knowledge in the system correspond to actuality?
- *Application*—to what extend does the system perform the intended task?

These are major dimensions in conventional systems analysis and in the evaluation of a scientific or technological system. The assumption in expert systems development is that these may be difficult, or impossible, criteria to apply because of the nature of the problem. For knowledge-based systems in general, however, these conventional standards are still very important.

The interactions of the system with some mental world, for example that of the clients, for whom it is intended to perform decision or control tasks introduce two dimensions:

- *Comprehension*—to what extent is the system able to understand information and queries?
- *Explanation*—to what extend is the system able to explain its activities?

These are important dimensions in any system supporting human activities. There have been attempts to satisfy them in most knowledge-based system development from MYCIN/TEIRESIAS onwards. Conventional data-processing has often been weak in these areas.

The interactions of the system with some knowledge world underlying its performance of decision or control tasks introduce two dimensions:

- *Derivation*—to what extent can the system derive its required knowledge from well-established sources?
- *Creation*—to what extend is the system able to generate new knowledge?

The rational reconstruction of knowledge to show, regardless of its source, how it could have been derived from well-established sources is important in expert systems development. We would prefer the minimal content to be dependent on the authority of experts as opposed to their use of professional knowledge. Similarly, while the creation of new knowledge is not the primary focus of system development, it is one of the most valued achievements—“we now think about the problem in different terms” is a significant positive comment resulting from a knowledge-based system development—it would be even more significant if it were a result of system activity.

There are three ancillary dimensions in Figure 8 relating to interactions between the worlds. Expert performance corresponds to a relation between world 2 and one of the others—the link shown is to world 1. We can evaluate the behavior of the system against the behavior of experts. This is a pragmatic evaluation based on the known utility of that expert behavior.

Expert knowledge corresponds to a relation between world 2 and world 3. We can evaluate the rationale of the system against the rationale of experts. This is a hermeneutic evaluation based on respect for the conceptual framework in which the experts perceive themselves as operating.

The objective knowledge of positive science is a relation between world 1 and world 2 that combines correspondence, derivation and explanation. It is usually not applicable in the situations for which expert systems are being developed. However, the concepts and techniques underlying knowledge-based systems in general should have a seamless interface to this highly refined form of knowledge.

9 Structures of Knowledge and their Technological Support

Knowledge-based systems have not developed in a vacuum. They are part of the evolution of computer technology and reflect trends in that technology towards end-user computing including end-user system development and maintenance. Knowledge is not something remote from human experience. It is closer to it than many of the technological devices that were necessary to bring computer technology into use in the early years of data processing. The use of knowledge, overtly and naturally expressed, to ‘program’ a computer system to provide well-presented and explained advice, is a natural step in computer product evolution.

There are many parallel paths converging on knowledge-based, user-oriented systems. *Hypertext* and *hypermedia* systems approach knowledge a different direction to that of expert systems, accepting it as having limited structure and processibility, and adding a limited structure through the links that is intended to aid the user rather than allow computer-based inferencing (Conklin 1986, Gaines & Vickers 1988). They extend knowledge representation to include any media whereby human knowledge can be transferred even if there is no means for analyzing that knowledge or its transfer. For example, a film of a skilled person performing a task which enables another person to perform that task better contains knowledge which can be accessed and transferred through a hypermedia system (Vickers & Kingston 1987).

Computer-aided software engineering (CASE), although the terminology emphasizes existing technological concepts, is itself a major step towards simplifying the application of computer systems (Martin & McClure 1988), and many of the concepts and techniques involved closely resemble those being developed for knowledge acquisition for knowledge-based systems (Gaines 1988c). A wide range of developments in computing, in scientific visualization, in animated simulation, in document preparation, in symbolic mathematics, in natural language interaction—all can be seen as *knowledge support* systems for knowledge-based human activities. In knowledge acquisition for knowledge-based systems we are often concerned to integrate these other developments. Do the models developed in this paper provide a framework for this integration?

Figure 9 analyzes major intellectual developments that are subject to technological support focusing on the path from human experience to the most refined scientific abstraction, currently category theory. It shows critical distinctions at each stage, for example, that between passive observation and the active, intentional processing of information that is peculiar to living systems. It shows the distinctions of Figures 4 and 5 leading through anticipation to modeling, and through interaction to control. It shows how experience is shared through communication allowing anticipation to be an activity of the species rather than of the individual. It shows how communication is refined through evocation of desired responses in rhetoric, accurate description of experience through representation, and the structured argument of discourse. It shows that the application of rationality criteria to discourse leads to the synthesis of new knowledge through induction and the analytic development of existing knowledge through deduction, arithmetic, logic, algebra and category theory.

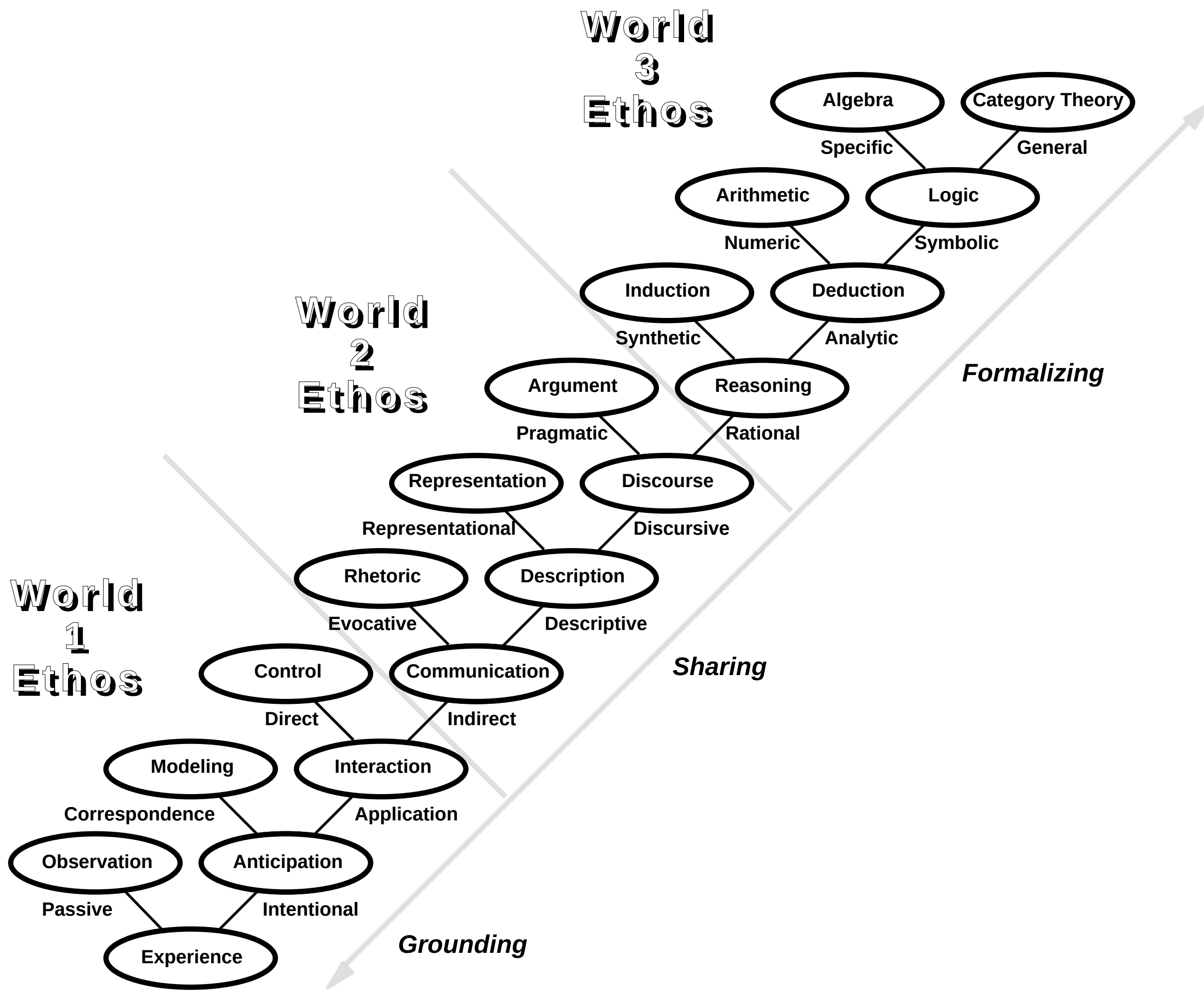

Figure 9 Human intellectual developments

Figure 9 is an extended example of the 'reflective equilibrium' between our experience and our models (Stich & Nisbett 1984)—we reject some aspects of experience that conflicts with our preferred models and we reject some aspects of our models in applying them to experience. The diagram shows that at each stage of formalization and definition additional presuppositions are introduced and some aspects of the preceding stage are lost. The path in Figure 9 appears long enough to raise questions as to the possibility of our most remote formalizations being applicable to our direct experience. However, there is an important meta-theorem relating to the 'application' of a finitely axiomatized mathematical theory—it is instantiable through the choice of constraints corresponding to its axioms in the way that we describe nature, or from the imposition of constraints corresponding to its axioms upon nature. That is, our cognitive perspectives and our technologies both create conditions in which our formalizations and definitions become applicable—they close the loop in Figure 9. In addition, if nature is truly chaotic it will generate an infinite variety of constraints sufficient to instantiate any finitely axiomatized theory through appropriate choice of constraints.

Figure 9 presents a level of detail below that Figure 7, marking some of the major intellectual developments of the human species in constructing an inductive hierarchy of knowledge. The table in Figure 10 sets out major knowledge processes from Figure 9, the relevant techniques and supporting technologies, pre- and post- the development of computing. Looking down the right hand column of Figure 10 we can see the full spectrum of computer-based techniques supporting human knowledge processes. It is interesting to note the way in which the development of computing has straddled the central region of discourse and inductive reasoning. Computing commenced with number crunchers, moved into simulation and instrumentation, and then into databases, graphics, text processing. The support for the explanatory processes of discourse and the argument forms of practical reasoning are still very weak, even though these underlie our colloquial use of the term ‘expertise.’ Current expert systems largely support deductive reasoning and may be seen as a natural extension of database technology to allow retrieval of logically derived fields.

I have shown ‘advice systems’ as a place marker for what might be expected to be the dominant role of an ‘expert system’ emulating the main activities of a human expert, the knowledge organization and creation activities of inductive, abductive and analogical reasoning that are an integral part of expertise in the model developed in this paper. These activities are treated as ‘updating’ and ‘maintenance’ in current expert system technologies, but little provision is made in the deductively oriented shells to support these activities through techniques discussed earlier such as anomaly detection and reporting. The current knowledge acquisition literature focuses on rapid prototyping tools for knowledge engineering rather than on the integration of acquisition activity with the ongoing application of the knowledge-based systems. This is reasonable at the current state of the art, but integration of acquisition and performance will become a major design issue as knowledge-based systems come into wider use and the problems and expense of ‘updating’ and ‘maintenance’ become apparent.

The need for integration of computing technologies is very apparent in Figure 10. Practical knowledge support systems must be capable of providing combinations of graphics, text processing, hypermedia, number crunching, databases, communications, theorem proving, databases, and semantic nets and expert system technologies. These are not highly separated functions but rather multiple aspects of human knowledge processing.

Human Knowledge Processes	*Relevant Techniques*	*Supporting Systems* *Pre-Computing*	*Post-Computing*
Experience	Perception, Memory	Instruments, Data Recording	Instrumentation, Data Storage
Anticipation	Cognition, Modeling	Modeling Techniques, Statistics	Data Analysis
Interaction	Action, Control	Tools, Mechanisms	Simulation, Robotics
Communication	Rhetoric, Mimicry	Pamphlets, Novels, Drama, Movies	Games
Description	Representation, Language	Representational Painting & Writing, Phonograph, Television	Word Processing, Graphics, Information Retrieval
Discourse	Argument, Communicative Action	Books, Journals, Journalism, Telephone, Education System	Electronic Mail, Hypermedia, CAL, Natural Language
Reasoning	Induction, Abduction, Analogy	Jurisprudence Systems, Uncodified 'Commonsense'	Semantic Nets, Conceptual Graphs, 'Advice Systems'
Deduction	Mathematics, Set Theory, Arithmetic	Calculators, Diagramming Techniques	Number Crunchers, Relational Databases, 'Expert Systems'
Logic	Propositional & Predicate Calculi, Modal Logics	Proof Techniques, Sequent Calculi	Algebraic Manipulators, Theorem Provers
Category Theory	Universal Algebras, Topology, Homology	Metatheorems, Proof Techniques	Category-Theoretic Programming Languages

Figure 10 Human Knowledge Processes, Techniques and Supporting Systems

10 Conclusions

A model of knowledge-acquisition for knowledge-based systems has been presented which presents the acquisition activity as playing an essential and continuous role in skilled performance, rather than as a separate and separable activity. The practical implications of this model for systems design have been developed, and recommendations made targeted on monitoring the quality of advice from expert systems and achieving closer integration between the application of these systems and the formation of expertise. The model has been developed in depth to generate taxonomies of human knowledge processes and use these to analyze the roles of a wide variety of computer-based systems in supporting these processes. The model has been used to highlight strengths and weaknesses in the current state of the art in knowledge-based systems. An overall framework has been provided for the variety of knowledge acquisition problems, techniques and technologies discussed in the literature.

Acknowledgements

Financial assistance for this work has been made available by the Natural Sciences and Engineering Research Council of Canada. I am grateful to many colleagues for discussions over the years that have influenced this paper. In particular I would like to thank John Boose, Jeff Bradshaw, Bruce Conrad, George Klir, Debbie Leishman, Gordon Pask, Mildred Shaw, Joan Vickers and Ian Witten, for access to their own research and many stimulating discussions. The final version was improved by critical comment from participants in the AAAI Knowledge Acquisition Workshop at Banff in November 1988. Much of the research reported in this paper is part of a long-term collaborative study with Mildred Shaw of the cognitive foundations of human knowledge processes.

References

Benbasat, I. & Dhaliwal, J.S. (1988). A framework for the validation of knowledge acquisition. Proceedings of the Third AAAI Knowledge Acquisition for Knowledge-Based Systems Workshop. pp.1-1-1-18. Banff (November).

Blume, S.S. (1974). Towards a Political Sociology of Science. New York: Free Press.

Boose, J.H. & Bradshaw, J.M. (1987) Expertise transfer and complex problems: using AQUINAS as a knowledge acquisition workbench for knowledge-based systems. International Journal of Man-Machine Studies(26), 3-28.

Conklin, J. (1986). A survey of hypertext. MCC Technical Report STP-356-86. Austin:Microelectronics & Computer Technology Corporation (October).

Davis, R. & Lenat, D.B. (1982). Knowledge-Based Systems in Artificial Intelligence. New York: McGraw-Hill.

Feyerabend, P. (1975). Against Method. London: NLB.

Gaines, B.R. (1987a) An overview of knowledge acquisition and transfer. International Journal of Man-Machine Studies 26(4), 453-472 (April).

Gaines, B.R. (1987b). How do experts acquire expertise? Proceedings of the First European Workshop on Knowledge Acquisition for Knowledge-Based Systems. pp.B2-1-11. UK: Reading University (September).

Gaines, B.R. (1987) Advanced expert system support environments. Boose, J.H. & Gaines, B.R. (Eds) Proceedings of the Second AAAI Knowledge Acquisition for Knowledge-Based Systems Workshop. pp.8-0-8-14. Banff (October).

Gaines, B.R. (1988a). Second generation knowledge acquisition systems. Proceedings of the Second European Workshop on Knowledge Acquisition for Knowledge-Based Systems (EKAW'88). GMD-Studien Nr.143. pp.17-1-17-14. Bonn, Germany: Gellsellschaft für Mathematik und Datenverarbeitung (June).

Gaines, B.R. (1988b). Positive feedback processes underlying the formation of expertise. IEEE Transactions on Systems, Man & Cybernetics, to appear.

Gaines, B.R. (1988c). Software engineering for knowledge-based systems. Proceedings of CASE'88, Second International Workshop on Computer-Aided Software Engineering. pp. 14-1-14-7.

Gaines, B.R. & Shaw, M.L.G. (1986). A learning model for forecasting the future of information technology. Future Computing Systems, 1(1), 31-69.

Gaines, B.R. & Vickers, J.N. (1988). Design considerations for hypermedia systems. Microcomputers for Information Management, 5(1), 1-27 (March).

Habermas, J. (1981). The Theory of Communicative Action: Reason and the Rationalization of Society. Boston: Beacon Press.

Hagstrom, W.O. (1965). The Scientific Community. New York: Basic Books.

Hawkins, D. (1983). An analysis of expert thinking. International Journal of Man-Machine Studies, 18(1), 1-47 (January).

Hayes-Roth, F. (1984). The industrialization of knowledge engineering. Reitman, W., Ed. Artificial Intelligence Applications for Business. pp. 159-177. Norwood, New Jersey: Ablex.

Klir, G.J. (1985). Architecture of Systems Problem Solving. New York: Plenum Press.

Martin, J. & McClure, C. (1988). Structured Techniques: The Basis for CASE. Englewood Cliffs, New Jersey: Prentice Hall.

Merton, R.K. (1973). The Sociology of Science: Theoretical and Empirical Investigations. Chicago: University Press.

Notturno, M.A. (1985). Objectivity, Rationality and the Third Realm: Justification and the Grounds of Psychologism: A Study of Frege and Popper. Dordrecht, Holland: Martinus Nijhoff.

Pask, G. (1980). Developments in conversation theory—Part I. International Journal of Man-Machine Studies, 13(4), 357-411 (November).

Popper, K.R. (1968). Epistemology without a knowing subject. Van Rootselaar, B., Ed. Logic, Methodology and Philosophy of Science III. pp. 333-373. Amsterdam, Holland: North-Holland Publishing Co.

Rappaport, A. & Gaines, B.R. (1988). Integration of acquisition and performance systems. Proceedings of the Third AAAI Knowledge Acquisition for Knowledge-Based Systems Workshop. pp.25-1-25-20. Banff (November).

Schön, D.A. (1983). The Reflective Practitioner. New York: Basic Books.

Shaw, M.L.G. & Gaines, B.R. (1988). A methodology for recognizing consensus, correspondence, conflict and contrast in a knowledge acquisition system. Boose, J.H. & Gaines, B.R. (Eds) Proceedings of the Third AAAI Knowledge Acquisition for Knowledge-Based Systems Workshop. pp.30-1-30-19. Banff (November).

Shaw, M.L.G. & Woodward, J.B. (1987) Validation of a knowledge support system. Boose, J.H. & Gaines, B.R. (Eds) Proceedings of the Second AAAI Knowledge Acquisition for Knowledge-Based Systems Workshop. pp.18-0-18-15. Banff (October).

Stich, S.P. & Nisbett, R.E. (1984). Expertise, justification and the psychology of inductive reasoning. Haskell, T.L., Ed. The Authority of Experts. pp. 226-241. Bloomington, Indiana: Indiana University Press.

Vickers, J.N. & Kingston, G.E. (1987). Modelling the master coach: building an expert system for coaching. Norrie, D., Ed. Learning in Future Education: Proceedings of International Conference on Computer Assisted Learning in Post-Secondary Education. pp.207-212. University of Calgary.