Strassoldo R 1980 Centre-periphery and system boundary: Cultural perspectives. In: Gottmann J (ed.) *Centre and Periphery. Spatial Variation in Politics*. Sage, Beverly Hills, CA,, pp. 27–61

Thrift N 1985 Flies and germs: A geography of knowledge. In: Gregory D, Urry J (eds.) *Social Relations and Spatial Structures*. Macmillan, Basingstoke, Hampshire, pp. 366–403

Törnqvist G 1968 *Flows of Information and the Location of Economic Activities*. Gieerup, Lund

World Bank 1999 *World Development Report 1998/99. Knowledge for Development*. Oxford University Press, Oxford, UK

P. Meusburger

Knowledge (Explicit and Implicit): Philosophical Aspects

Explicit knowledge is knowledge that the knower can make explicit by means of a verbal statement: 'Someone has explicit knowledge of something if a statement of it can be elicited from him by suitable enquiry or prompting' (Dummett 1991). Implicit knowledge can then be defined simply as knowledge that is not explicit. On this construal, implicit knowledge corresponds roughly to what Polanyi calls 'tacit knowing': 'we can know more than we can tell' (1967).

The distinction between explicit and implicit knowledge might seem to be connected closely to the familiar distinction (Ryle 1949) between knowledge of a proposition ('knowing that') and possession of a skill ('knowing how'). In everyday cases of 'knowing that' (such as knowing that the atomic number of lead is 82) the knower can say what it is that they know. In contrast, it is manifestly possible to have a skill (such as the ability to tie one's shoelaces) while being quite unable to give any verbal account of how the performance is achieved. So 'knowing how' may provide examples of Polanyi's 'tacit knowing.'

But implicit knowledge, as it is considered here, is closer to 'knowing that' than to 'knowing how.' In cases of implicit knowledge, a proposition or rule is known, but this is not available to the knower for verbal report. In the absence of verbal report, an attribution of implicit knowledge must be supported by other kinds of empirical evidence. But more fundamentally, we need some account of what it is for a subject to possess knowledge but to be quite unable to make it explicit (cf. Dienes and Perner 1999).

1. Implicit Memory and Perception

In cognitive psychology, a distinction is drawn between explicit and implicit tests of memory. Explicit memory tests 'make explicit reference to and demand conscious recollection of a specific previous experience.' But in an implicit memory test 'memory for a recent experience is inferred from facilitations of performance … that need not and frequently do not involve any conscious recollection of the prior experience' (Schacter 1989). For example, a subject might be shown a list of words and then tested in various ways. In an explicit test, the subject might be asked to say which words were on the list (recall) or might be presented with words and asked whether they appeared on the list (recognition). In an implicit test, the subject might be given a task that involves responding to words (e.g., by reading them aloud). Memory for the words on the list would be inferred from the fact that the subject reacts faster to those words than to words that were not on the list. The subject's reaction to a word that was on the list is said to be 'primed' by the earlier exposure to that word.

Intuitively, explicit tests of recall or recognition probe whether the subject still knows, in the everyday sense, which words were on the list. In contrast, implicit tests seem to provide evidence that information about which words were on the list is in some way present in the subject, perhaps in a form that can influence the subject's performance only on particular tasks. This intuitive idea that the two kinds of tests probe rather different kinds of memory, or information storage, is supported by findings that there are factors that affect performance on explicit tests and implicit tests differently (Schacter 1989). For example, extending the interval between the initial exposure to the list and the subsequent test may impair performance on an explicit test but not on an implicit test. Conversely, changing the sensory modality of presentation between initial exposure to the list and subsequent test (e.g., from auditory to visual presentation) significantly reduces priming effects in an implicit test but has little or no effect on recall and recognition in an explicit test. In addition, under some conditions performance on an implicit test has been shown to be uncorrelated with success or failure on an explicit test.

Priming effects may be observed even in the total absence of recall or recognition. When a word is initially presented with a very short exposure, and is followed by another stimulus to produce 'backward masking', the subject is quite unable to report anything about the word. But that initial presentation may still facilitate responses to the same word in a subsequent task. The results of masked priming experiments demonstrate implicit memory without explicit memory.

In cases where a subject is unable, even at the time of the first presentation, to report which word appears on the screen, we may say that perception itself is implicit rather than explicit: it is 'perception without awareness.' The influence of perceptually presented information of which the subject is unaware is also apparent in an experiment involving very brief presentations of photographs of faces that show expressions of emotions (Dimberg et al. 2000). Subjects are shown a happy, neutral, or angry face, rapidly followed by a

neutral face (to produce backward masking). They report seeing nothing before the neutral face. Yet their facial reactions, detected by measuring activity in facial muscles, are significantly different, depending on the emotional expression of the briefly presented face.

2. Knowledge and Belief

It is natural to describe these cases of implicit memory and implicit perception by saying that the tests reveal that the subject knows more than she can tell. But it is also natural to raise a query about the use of the term 'know' here. In epistemology, knowledge is usually taken to be true belief that meets some further condition involving justification of the belief or reliability of the method by which the belief is produced. The mechanisms of information gathering in perception and information storage in memory are plausibly reliable. But the notion of justification does not seem to be applicable in cases where the subject is unaware of the presence or influence of the information. So, if knowledge always involves justification, then so-called implicit knowledge does not seem to meet the requirements for knowledge. And even if knowledge is construed in a more inclusive way, there is room for doubt as to whether implicit knowledge amounts to a belief on the part of the subject.

At least three factors contribute to this doubt. One is the subject's lack of awareness. But this is not, by itself, decisive since it is possible to make sense of the notion of an unconscious belief. A second factor is that implicit knowledge seems to be limited in its influence. A belief may figure in many different ways in the subject's subsequent reasoning and planning, depending on their other beliefs and desires. But implicit knowledge that a happy face has been presented has only been shown to affect facial muscles. States of implicit knowledge do not appear to be inferentially integrated with the subject's beliefs (Stich 1978); they are, rather, inferentially isolated. However, this again is not quite decisive, since we can make sense of the idea of compartmentalized beliefs.

A third factor contributing to the doubt about the applicability of the folk psychological notion of belief is that believing requires the possession of concepts. No one can believe that penguins waddle without having some concept of penguins and of waddling. But, in at least some of the cases of implicit knowledge that we have mentioned, the subject does not need to possess the concepts that a theorist would use to characterize the information that has been gathered and stored. Thus, for example, the involuntary facial response to a happy or angry face does not seem to depend on the subject having any conceptualized grasp of the emotions of happiness or anger. A subject who lacks the concepts of happiness or anger cannot have any beliefs about happiness or anger. Yet such a subject may, it seems, be described as possessing implicit knowledge that a happy face or an angry face has been presented.

These considerations do not show that cases of implicit knowledge should never be grouped together with ordinary knowledge and belief. But examples of implicit memory and perception illustrate how, in cognitive psychology, the notion of implicit or tacit knowledge encompasses unconscious storage and processing of information whether or not this falls within the domain of the folk psychological notion of belief.

3. Preserved Implicit Knowledge in Neuropsychological Patients

Patients suffering from brain damage as a result of a stroke or head injury often show preserved implicit knowledge in the absence of explicit knowledge (Schacter et al. 1988).

Amnesic patients cannot remember recent experiences. If, for example, an amnesic patient is given a list of words to study she will perform poorly on a subsequent test of recall or recognition. But studying the list, even though it goes unremembered, may affect subsequent performance. In some cases, amnesic patients who are asked to complete three-letter sequences with the first words that come to mind, show the influence of the studied list in much the same way as happens with normal subjects (Warrington and Weiskrantz 1968). This is implicit memory in the absence of explicit memory.

Some patients are blind in a region of their visual field as a result of damage to their visual cortex. They are unable to report whether a light stimulus is present or absent in that region, to say whether a presented stimulus is, for example, an X or an O, or to discern the emotional expression on a face presented in the blind field. The patients say that they do not see the stimulus at all. Yet some of these patients, when asked to guess whether a stimulus is present or absent, or whether the stimulus is an X or an O, or whether a video clip presented in the blind field shows a happy or a fearful face, perform significantly above chance levels. This is the phenomenon of 'blindsight' (Weiskrantz 1986, de Gelder et al. 1999). It is a case of implicit perception in the absence of conscious or explicit awareness of the salient properties of the presented stimulus.

Prosopagnosic patients are unable to recognize and identify familiar faces. If, for example, a patient is asked to classify photographs as being of familiar or unfamiliar faces or to match names to faces, they perform at chance levels. Yet their autonomic affective responses (in particular, skin conductance responses) to familiar faces are significantly different from those to unfamiliar faces. A prosopagnosic patient may be unable to classify the faces of famous people according to their occupation (e.g., politician or television personality) although they can, of course, correctly assign occupations to the names of these people. But,

in some cases, presenting the face of a television personality alongside the name of a politician or vice versa interferes with the patient's performance when asked to assign an occupation to the name, just as it does in normal subjects (Young 1998). Information about the occupation of the person whose face is presented affects performance even though it is not available for verbal report.

Blindsight and other neuropsychological phenomena present challenges for both philosophical and empirical theories about consciousness. They also raise questions about the extent to which normal subjects are authoritative about the workings of their own minds. A normal subject would probably find it compelling to suppose that, when he sees the face of a television personality, his conscious recognition of the face interferes with his classification of a simultaneously presented name as that of a politician. Although this is likely to be correct, the fact that the same pattern of interference is found in patients who are unable to recognize faces raises an alternative possibility. It is at least conceivable that, contrary to what it is so compelling to suppose, the interference is produced in normal subjects in the same way as in prosopagnosic patients, by a process of which the subject is quite unaware.

4. Implicit Learning

Subjects may be trained to classify items as belonging or not belonging to some category by being provided with a principle or rule of classification. That is, the teacher may make the principle or rule explicit. Alternatively, subjects may be provided with a set of exemplars of the category; or they may simply be provided with feedback as to whether their classifications are correct or incorrect. On the basis of studying a set of exemplars, or after a period of training, a subject may achieve performance on the classification task that is significantly better than guessing at random. Indeed, the subject's classifications may be reliably correct. A subject who has attained a reliable level of performance may then be able to state the principle or rule of classification. In this case, even without being explicitly told the rule, the subject has gained explicit knowledge of the rule. But it may happen that the subject is quite unable to state the rule, even though his classifications reliably conform to the rule. In this case, the subject may be said to have learned the rule implicitly. The subject has achieved conformity to the rule without being able to state the rule.

Research in the area of implicit learning involves many complex issues (Berry and Dienes 1993, Shanks and St. John 1994). An experimenter may assign items to a category on the basis of a particular rule of classification, R, and may seek to discover whether a subject who has learned to make the same category assignment has arrived at explicit knowledge of the same rule. But asking the subject whether he is using rule R to make his classifications is problematic since considering that question may itself bring about a change in the subject's explicit knowledge.

Instead, the subject may simply be asked what rule they are using. But, if they do not produce a statement of rule R in response to this question, it by no means follows that the subject's classifications depend on implicit rather than explicit knowledge. One reason is that a rule that is explicitly known might not be volunteered because, for example, the subject is not entirely confident that it is correct. A second reason applies even if the subject does not have explicit knowledge of rule R. It may be possible to achieve classification performance that is significantly better than guessing at random, or even reliably correct over a large range of cases, by using explicit knowledge of some other rule.

To at least some extent, problems with the notion of implicit learning can be overcome by identifying independently well-motivated characteristics of implicit and of explicit knowledge. If a subject does not provide a statement of rule R but nevertheless makes classifications in conformity with rule R then their classification performance can be assessed in the light of these characteristics.

There is, however, a more general problem with the notion of implicit learning of a rule. So far, this has been characterized as learning to make classifications that conform to the rule, yet without explicit knowledge of the rule. We have observed some difficulties in deciding empirically whether explicit knowledge is really absent. But the more general problem is that any battery of performance that conforms to a rule inevitably conforms to more than one rule. If a subject's classification performance conforms to rule R_1 and to rule R_2 and the subject has explicit knowledge of neither rule, then the subject may be said to have implicitly learned something. But that does not settle whether the attained state of the subject should be described as implicit knowledge of R_1 or of R_2. In a case where two rules require just the same classifications, there is a question of principle whether any sense can be given to the notion of implicit knowledge of one rule rather than the other. A considerable body of research on implicit learning concerns the task of learning an artificial grammar (Reber 1989). But it is in the context of discussions of knowledge of natural language grammars that this question of principle has its original home.

5. Tacit Knowledge of Rules

In a famous passage, Chomsky (1965) says: 'Obviously, every speaker of a language has mastered and internalized a generative grammar that expresses his knowledge of his language. This is not to say that he is aware of the rules of the grammar or even that he can become aware of them.' Knowledge of the rules or

principles of a generative grammar is usually described as tacit knowledge and it counts as implicit knowledge according to our definition. The claim that ordinary speakers of natural languages tacitly know a grammar for their language is a core element of Chomskyan linguistics (e.g., Chomsky 1986, 1995).

Some philosophers have argued that the claim involves a conceptual confusion on the grounds that the notion of a rule of language belongs with the idea of a normative practice in which people advert to rules to justify, criticize or excuse their actions (Baker and Hacker 1984). But the notion of a tacitly known rule should be distinguished from the notion of a rule that figures in a normative practice. Once the distinction is made, there remains an important question about the relationship between the rules that are tacitly known by individual speakers and the normative practices in which those speakers participate. But that question does not indicate any conceptual incoherence in the very idea of tacit knowledge.

Other critics have urged that the ideas of tacit knowledge of rules and of unconscious information processing in general are problematic because of a deep connection between intentionality and consciousness (Searle 1990). But in response to this worry it can be said that, while there may be a connection between consciousness and the intentionality of beliefs, the states of tacit knowledge that Chomsky talks about are importantly different from belief states (Stich 1978, Davies 1989).

5.1 *Quine's Challenge*

Quine challenges Chomsky's introduction of the notion of tacit knowledge by making use of the distinction between behavior that conforms to a rule and behavior that is guided by a rule. A subject can behave in a way that conforms to a rule without using the rule to guide his behavior for, as Quine (1972) uses the notion of guidance: '(T)he behavior is not *guided* by the rule unless the behaver knows the rule and can state it.' Guidance requires explicit knowledge.

Chomsky's tacit knowledge is supposed to require less than explicit knowledge; but it cannot be equated with mere conformity. In fact, conformity to rules is neither necessary nor sufficient for tacit knowledge of those rules. It is not necessary, since the presence of tacit knowledge of rules does not guarantee perfect deployment of that knowledge in actual performance. It is not sufficient, since a tacit knowledge claim is not offered as a summary description of behavior but as a putative explanation of behavior. There will always be alternative sets of rules that require just the same behavior for conformity; but it is part of the idea of tacit knowledge that a speaker's actual behavior might be correctly explained in terms of tacit knowledge of one set of rules rather than the alternatives.

It is at this point that Quine (1972) poses his challenge. He insists that, if an attribution of tacit knowledge is an empirical claim that goes beyond a summary of conforming behavior, then it should be possible to indicate what kinds of evidence would count in favor of or against that empirical claim. He also insists that this evidence should involve the subject's behavior. To this latter point, it is reasonable to reply that there can be no a priori limit on the kinds of evidence that might be relevant to an empirical claim. So it is not legitimate to restrict evidence to the behavior of the very subject to whom the attribution of tacit knowledge is being made. But the more general point about evidence is a fair one.

5.2 *Sketch of an Account of Tacit Knowledge*

More fundamental, however, than the question of what evidence would support an attribution of tacit knowledge is the question what the correctness of such an attribution would consist in. We can sketch one answer to this fundamental question by using an example that involves not rules of syntax but very simple letter-sound rules of the kind that could be employed in reading words aloud. Suppose that one of these rules states that if a word begins with the letter 'b' then its pronunciation begins with the sound /B/. If a subject's pronunciation behavior conforms to this rule then it displays a pattern. Whenever a presented word begins with 'b,' the subject's pronunciation begins with /B/. But the transitions that concern us are not these transitions from presentation to pronunciation. Rather, we focus on transitions amongst beliefs or states of information. Whenever the subject starts out with the information that the presented word begins with 'b,' the subject ends up with the information that the word's pronunciation begins with /B/.

If these states were beliefs, then the subject's pattern of transitions from belief state to belief state would be accounted for if the subject possessed explicit knowledge of the 'b'-/B/ rule. This piece of explicit knowledge would figure as a common factor in the causal explanations of the subject's transitions from belief to belief. In contrast, there would be no such common factor if the subject had merely memorized the pronunciation of each of a large number of words beginning with 'b.' The difference between having explicit knowledge of the rule and having an independent piece of knowledge for each of the instances that fall under the rule corresponds to a difference in causal-explanatory structure.

In fact, we do not assume that the subject has beliefs about words and their pronunciations; the transitions may involve states of the kind that figure in information-processing psychology. Still less do we assume that the subject either has explicit knowledge of pronunciation rules or else has explicitly memorized the pronunciations of words. But we do make use of the idea of a common factor in the causal explanations of transitions. An attribution of tacit knowledge of the

'b'-/B/ rule can be construed as the claim that there is a single state of the subject that figures in the causal explanations of the various particular transitions that the subject makes (Evans 1981).

In general, a state of tacit knowledge is a state that figures as a common factor in causal explanations of certain transitions amongst states of information (or beliefs). Where different sets of rules require the same transitions for conformity, the attribution to a subject of tacit knowledge of a particular set of rules is made correct by the presence of a particular structure in the causal explanations of the subject's transitions (see also Davies 1987, Peacocke 1989).

Once an account of tacit knowledge in terms of causal-explanatory structure has been given, it is a relatively straightforward matter to give examples of empirical evidence that would confirm the attribution to a subject of tacit knowledge of a particular set of rules such as a grammar. Indeed, some of this evidence meets Quine's additional requirement of concerning the behavior of the subject to whom the attribution is being made. Relevant behavioral evidence could come from experimental studies of language acquisition, language perception, and language dysfunction following brain damage, while further evidence would be available from neural imaging.

6. Explicit and Implicit Representation

According to the basic notion of explicit knowledge introduced at the beginning of this article, explicitness is a matter of the subject being able to present information in linguistic form. It is not a matter of how the information is stored in the subject's mind or brain when it is not being called upon.

Suppose that someone knows, in the everyday sense of that term, the axioms of some theory. Provided that this knowledge can be verbally articulated it counts as explicit knowledge. Now consider some theorem that is derivable from those axioms. The person who knows the axioms may well, with some suitable enquiry and prompting, be able to see that the theorem follows from the axioms and to state it verbally. For example, someone who explicitly knows some elementary arithmetical facts may be able to work out, and to state, that 68 plus 57 equals 125. On Dummett's account, this latter piece of knowledge, even though it is computed when needed rather than being stored in memory, counts as no less explicit than the stored elementary facts from which it is derived. Indeed, it counts as explicit knowledge even before the knower works it out; for explicitness is defined in terms of the possibility of eliciting a verbal statement by enquiry or prompting. Once some propositions are classified as explicitly known, the category of explicit knowledge also includes at least some of the as-yet-undrawn consequences of those propositions.

But the terms 'explicit' and 'implicit' are also used to mark a number of distinctions that are specifically related to the storage and processing of information. Thus, Dennett introduces a quite different distinction from Dummett's when he says (1983): 'Let us have it that for information to be represented *implicitly*, we shall mean that it is *implied* logically by something that is stored explicitly.' On this different usage, the as-yet-undrawn consequences of propositions that count as explicitly stored would be classified as implicitly represented.

In fact, relative to any given notion of explicit representation or storage of information, it is possible to define a whole family of notions of implicit representation or knowledge. The members of the family differ over the inferential resources that can be used in drawing out consequences from the information that is explicitly stored. We might, for example, consider only deductive inferences that can be carried out by the person in question. Then, if a person explicitly stores elementary arithmetical facts in memory and is able to infer that $68 + 57 = 125$, this would count as implicit representation of that latter fact. Alternatively, we might allow any valid deductive inference, whether or not the person can actually carry it out. In that case, implicit representation or knowledge would outrun the possibility of eliciting from the person a verbal statement. We might go beyond deductive logic and allow methods of induction or rule extraction, with or without a restriction to methods that are available to the person in question. In that case, a person who stores in memory only facts about particulars might count as having implicit knowledge of generalizations or rules that subsume those facts.

The notion of implicit representation also depends on a prior conception of explicit storage. But there is no uniquely correct account of what explicit storage amounts to. One idea that is, however, fairly naturally associated with explicitness in the context of theories of information storage and processing is that explicitly stored information has to be accessed before it can be used. An access problem may result in a failure to use stored information that is relevant to the task in hand. A second idea that sometimes goes along with the notion of explicitness is that explicit storage involves the use of a language-like code. Putting the two ideas together, we have a picture in which explicit storage of information is a matter of linguistic inscriptions on pages that are not always easy to locate. Drawing out the consequences of explicitly stored information would then be naturally conceived as manipulating inscriptions in ways that are analogous to the formal procedures involved in the derivation of theorems from axioms in logical systems (Fodor 1987).

The notions of explicit storage and implicit representation under discussion in this section belong with mechanistic theories about the storage and processing of information. They could be applied to any information-processing systems and have no

special connection with human subjects. Even in the case of human subjects, information might be explicitly stored yet not accessible to consciousness and not available to the subject for verbal report.

A state of tacit knowledge (as sketched in Section 5) could be constituted by the explicit storage of a rule that was not accessible to consciousness. Indeed, it sometimes seems to be assumed that, within the picture of the storage and processing that we have just sketched, tacit knowledge of a rule is always embodied in a mechanism by having the rule explicitly stored. There is a clear distinction between explicitly storing information and using that information; use of stored information involves processes for accessing that information. But such access processes are not typical of the operation of connectionist networks (McClelland et al. 1986): 'Using knowledge in processing is no longer a matter of finding the relevant information in memory and bringing it to bear; it is part and parcel of the processing itself.' So, against the background of the assumption that tacit knowledge involves explicit storage, it is often said that what is distinctive of connectionist networks is that they perform cognitive tasks without having tacit knowledge of rules.

The account of tacit knowledge in terms of causal-explanatory structure certainly allows for the case of tacit knowledge that is a matter of explicit storage. But the key notion in the account is that of a common factor in causal explanations of certain transitions amongst states of information. So the account also allows for the possibility of tacit knowledge that is 'part and parcel of the processing,' and it is left open that tacit knowledge may sometimes be embodied in connectionist networks (Davies 1995).

See also: Cognitive Science: Overview; Implicit Learning and Memory: Psychological and Neural Aspects; Intentionality and Rationality: A Continental-European Perspective; Intentionality and Rationality: An Analytic Perspective; Meaning and Rule-following: Philosophical Aspects; Perception: Philosophical Aspects; Reference and Representation: Philosophical Aspects

Bibliography

Baker G P, Hacker P M S 1984 *Language, Sense and Nonsense: A Critical Investigation into Modern Theories of Language*. Blackwell, Oxford, UK

Berry D C, Dienes Z 1993 *Implicit Learning: Theoretical and Empirical Issues*. Lawrence Erlbaum Associates, Hove, East Sussex, UK

Chomsky N 1965 *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA

Chomsky N 1986 *Knowledge of Language: Its Nature, Origin, and Use*. Praeger, New York

Chomsky N 1995 *The Minimalist Program*. MIT Press, Cambridge, MA

Davies M 1987 Tacit knowledge and semantic theory: Can a five per cent difference matter? *Mind* **96**: 441–62

Davies M 1989 Tacit knowledge and subdoxastic states. In: George A (ed.) *Reflections on Chomsky*. Blackwell, Oxford, UK

Davies M 1995 Two notions of implicit rules. In: Tomberlin J E (ed.) *Philosophical Perspectives, 9: AI, Connectionism, and Philosophical Psychology*. Ridgeview, Atascadero, CA

de Gelder B, Vroomen J, Pourtois G, Weiskrantz L 1999 Non-conscious recognition of affect in the absence of striate cortex. *NeuroReport* **10**: 3759–63

Dennett D C 1983 Styles of mental representation. *Proceedings of the Aristotelian Society* **83**: 213–26. Reprinted in: *The Intentional Stance*. MIT Press, Cambridge, MA, 1987.

Dienes Z, Perner J 1999 A theory of implicit and explicit knowledge. *Behavioral and Brain Sciences* **22**: 735–808

Dimberg U, Thunberg M, Elmehed K 2000 Unconscious facial reactions to emotional facial expressions. *Psychological Science* **11**: 86–9

Dummett M 1991 *The Logical Basis of Metaphysics*. Harvard University Press, Cambridge, MA

Evans G 1981 Semantic theory and tacit knowledge. In: Holtzmann S H, Leich C M (eds.) *Wittgenstein: To Follow a Rule*. Routledge and Kegan Paul, London. [Reprinted In: *Collected Papers*. Oxford University Press, Oxford, 1985]

Fodor J A 1987 Why there still has to be a language of thought. In: *Psychosemantics*. MIT Press, Cambridge, MA

McClelland J L, Rumelhart D E, Hinton G E 1986 The appeal of parallel distributed processing. In: Rumelhart D E, McClelland J L, the PDP Research Group (eds.) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition Volume 1: Foundations*. MIT Press Cambridge, MA

Peacocke C 1989 When is a grammar psychologically real? In: George A (ed.) *Reflections on Chomsky*. Blackwell, Oxford, UK

Polanyi M 1967 *The Tacit Dimension*. Routledge and Kegan Paul, London

Quine W V O 1972 Methodological reflections on current linguistic theory. In: Davidson D, Harman G (eds.) *Semantics of Natural Language*. Reidel, Dordrecht

Reber A S 1989 Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General* **118**: 219–35

Ryle G 1949 *The Concept of Mind*. Hutchinson's University Library, London

Schacter D L 1989 Memory. In: Posner M I (ed.) *Foundations of Cognitive Science*. MIT Press, Cambridge, MA

Schacter D L, McAndrews M P, Moscovitch M 1988 Access to consciousness: Dissociations between implicit and explicit knowledge in neuropsychological syndromes. In: Weiskrantz L (ed.) *Thought Without Language*. Oxford University Press, Oxford, UK

Searle J 1990 Consciousness, explanatory inversion, and cognitive science. *Behavioral and Brain Sciences* **13**: 585–95

Shanks D R, St. John M F 1994 Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences* **17**: 367–95

Stich S 1978 Beliefs and subdoxastic states. *Philosophy of Science* **45**: 499–518

Warrington E K, Weiskrantz L 1968 New method of testing long term retention with special reference to amnesic patients. *Nature* **217**: 972–4

Weiskrantz L 1986 *Blindsight: A Case Study and Implications*. Oxford University Press, Oxford, UK
Young A W 1998 *Face and Mind*. Oxford University Press, Oxford, UK

M. Davies

Knowledge Representation

1. Knowledge Representation: Ideal and Reality

The fundamental conjecture of knowledge-based artificial intelligence (AI), known as the physical symbol system hypothesis (Newell and Simon 1976), posits that the cognitive processes of any broadly intelligent agent, natural or artificial, can be largely characterized in terms of a knowledge base: a transparent, symbolic representation of the agent's beliefs, intentions, and value judgments, together with reasoning processes that modify the knowledge base in rational ways, and perceptual and motor processes that connect the knowledge base to the outside world. The field of *knowledge representation* (KR) is the study of how such a knowledge base can be constructed.

The physical symbol system hypothesis requires that a knowledge base be symbolic, transparent, and modular. That is, the atomic elements of a knowledge base are symbols, each of which has an associated meaning. It is possible to identify small substructures of the knowledge base that correspond to specific beliefs, judgments, and intentions. For example, the belief that Holbein painted a portrait of Thomas More in oil might be represented in a predicate calculus formula like

$$\exists_X \text{ oil\_painting}(X)$$
$$\wedge \text{ painter(holbein, } X) \wedge \text{ depicts } (X, \text{ thomas\_more})$$

or in a fragment of a semantic net, such as that shown in Fig. 1. Here the symbols are '$\exists$' (meaning 'there exists'), 'X' (a variable), '$\wedge$' (meaning 'and'), 'oil_painting,' 'painter,' 'holbein,' and 'thomas_more'; the parentheses and the comma indicate the syntactic structure. (The fact that some of these symbols are English words is purely for the convenience of the human reader; the formula could just as well be notated

$$\exists_X \text{ g001}(X) \wedge \text{ g002(g003, } X), \wedge \text{ g004}(X, \text{ g005})$$

if it is accepted that 'g001' denotes the property of being an oil painting, that 'g002' denotes the relation of a person painting a painting, and so on.) It should be noted that not every belief of the agent can be explicitly expressed in the knowledge base. For example, a person who knows that Holbein painted a portrait of Thomas More will also know that Holbein knew how to paint; that Thomas More was painted; that Thomas More was born before Holbein's death; and so on. Clearly it is not plausible that each of these ancillary beliefs is separately represented in the knowledge base. Rather, they are represented implicitly, in the sense that they can be derived as needed.

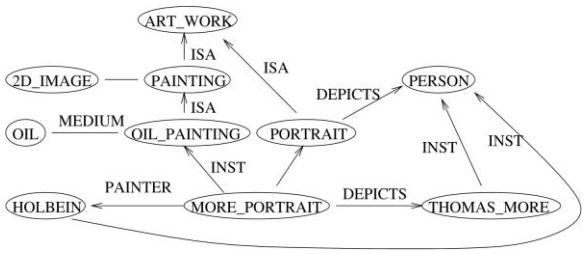

Figure 1
Fragment of a semantic net

Of course, the construction of the broadly intelligent artificial agent described above is today just a pipe dream for the distant future. Any work aimed at achieving a working system any time soon must be much more modest in its ambitions. Within the context of such limited objectives, the problem of knowledge representation is one of finding data structures that express the knowledge needed for the particular application.

Thus broadly construed, the problem of knowledge representation would appear to be equivalent to computer programming generally. Any program, after all, uses some data structures, and incorporates some kind of knowledge. What distinguishes a data structure as interesting from the point of view of knowledge representation is that its design reflects some of the features of the broadly intelligent knowledge-based agent described above. As artificial intelligence is the study of tasks that are easy for people (or animals) to carry out but difficult to implement in a computer, so knowledge representation is the design of data structures that express knowledge that people deal with easily, but are hard to encode in a data structure.

The design of a knowledge representation can be very roughly divided into four parts

(a) *Architectural* issues involve defining general-purpose schemas for knowledge representation and procedures for reasoning with these schemas. Most of the research in knowledge representation has been addressed to architectural issues, as they are largely independent of the particular domain and application (Reichgelt 1991).

(b) *Content* issues address such questions as what is the knowledge to be represented, how should the domain be conceptualized, what are the key concepts in the domain, and what kinds of partial knowledge must be dealt with. Since content issues are always

International Encyclopedia of the Social & Behavioral Sciences ISBN: 0-08-043076-7