Computational Intelligence, Volume 18, Number 1, 2002

IMPLICIT VERSUS EXPLICIT REPRESENTATION AND INTRA- VERSUS INTER-MODULAR PROCESSING

JOSEF PERNER
University of Salzburg, Austria

ZOLTAN DIENES
University of Sussex, England

Frawley raises a very interesting issue: Can we draw on lessons learnt in computer science about different types of control processes, and the distinction between ''control'' and ''logic'' (Kowalski 1979), to shed light on the human mind? Frawley concentrates discussion on unit-level control, i.e., the control of the flow of information across modules, and the representations (logic) that exist within modules. He classifies a group of developmental syndromes (e.g., Williams, autism) as suffering from an across module control problem in contrast to, e.g., SLI (specific language impaired) children, who suffer from a within module impairment of linguistic knowledge.

Frawley identifies one of the requirements of across module communication as being explicit, and finds our (Dienes and Perner 1999) attempt to explicate the meaning of the implicit explicit distinction particularly helpful. Frawley gives an admirably clear and succinct rendition of our account of predication and factuality explicitness, and uses it for his purposes. Frawley suggests that intermodular processing requires predication and factuality implicitness, whereas intramodular processing can do without: ''Representations that are within domains are preferentially implicit and procedural because they must apply generally (i.e., no explicit factivity). . . .By contrast, representations reported out of a domain have to be maximally explicit in order to be checked and used.'' We wish to elaborate on this theme.

Following our original article, we can illustrate predication implicitness-explicitness with the example of individual animals being classified as cats or dogs. If all that is generated is a label ''cat'' then although this represents the presented individual as a cat, it only makes the property of being a cat explicit and leaves it implicit that there is an individual of which this property of cat-ness is being predicated (predication implicit knowledge). On the other hand, the representation ''This is a cat,'' would make this predication explicit. A representation leaves its factuality implicit if it is simply taken as true by the system; it makes its factuality explicit if it can be represented as true or false or possibly true. Predication explicitness is necessary for tracking the same individual across different contexts. How might this be related to representations used within as opposed to between modules? Modules are defined to be ''informationally encapsulated'' (Fodor 1983); they operate only on certain types of information, they perform very specific operations on that information, and they do so online. By contrast, non-modular processing central processing, the global workspace (Baars 1988)—is inferentially promiscuous, dealing in principle with any type of content domain, making that information available to in principle any type of relevant processing, with in principle information coming in from any of a number of different modules about possibly different individuals. Given this idealised description, it would seem that information communicated by one module to other modules

Address correspondence to Josef Perner at the Institut fuer Psychologie, Universitaet Salzburg, Hellbrunnerstrasse 34, A-5020 Salzburg, Austria. E-mail: josef.perner@sbg.ac.at.

via central processing, could not usefully be predication implicit. If the representation is simply a feature, the system has no way of combining this representation with others to draw inferences; inferences require that propositions about the same individual can be taken to be about the same individual. An inferentially promiscuous system must be able to keep track of what individuals are being referred to in the different propositions it entertains. In contrast, a dedicated online module can take as given that the succession of operations performed refer to the immediately presented individual. So could the predication implicit-explicit distinction be a general one for intra- versus inter-modular communication or is this conclusion too quick? Imagine a processor is concerned about individual linguistic entities, e.g., a word or a sentence, which have a certain property, i.e., a certain linguistic structure. Language processing consists in this structure (property) of the presented item being transformed into different levels of description, e.g., from a phonetic specification to a specification in terms of lexical items embedded in a syntactic structure. Our above comments and Frawley's comments imply that within a module only the relevant structure of the linguistic item processed needs to be represented, but across modules there needs to be explicit representation of the entity being processed and that the represented linguistic structures belong to this particular entity. Let's be clear about why or under what conditions this might or might not be true. If we think of a module as transforming the structure from input to output without representing that this transformed structures belong to a particular entity, then why can't the next module just accept the output structure of its predecessors and work on it? Frawley does go on to say, "Σ in explicitly representing predication and factivity, declarative knowledge signals by its form its application to particular cases: Only in this explicit form can the representation be fully usable by another domain because its applicability and validity can be tracked." This need occurs when information is passed into a central workspace, as remarked above. But modules may not always communicate with each other via a central workspace, maybe the connection between some modules, e.g., morphology and phonology, is more direct. Then the need for predication explicitness would not occur simply because information is passed on from one to the next module. However, when one module needs to check on what another module is or has been doing while still working on the case, the need does arise. For instance if the morphology module needs some information from the phonology program about a detail that was not passed on automatically or could not be resolved on pure phonetic grounds without information from higher up, then, indeed, the interchange between modules must make explicit which processed entity the information pertains to, e.g., which phoneme or group of phonemes (explicit predication), if other phonemes are currently being processed by the phonology module. If, however, the system is constructed such that all modules likely to need exchange of information have to work on the same entity (group of phonemes) then predication can remain implicit since it is clear what entity (group of phonemes) the requested information is about. Predication explicitness is needed not only for tracking individuals across contexts, but also binding different features to different individuals. Thus, a vision module may need to use predication explicit representations within itself to complete a fully structured processing of the outside world.

In sum, we regard predication explicit representation to be favoured in inter-modular communication and predication implicit representation to be favoured in intra-modular communication, but this mapping need not be perfect.

Frawley also requires that across module communication be factuality explicit. Our idea behind factuality explicitness is that even when there is explicit predication (*This is a cat*) then, unless we are restricted to process only veridical representations about the world, we also need to distinguish whether this proposition holds in the real world or in some hypothetical or counterfactual world. Now, online language parsing across all

involved modules, one would think, is restricted to just processing information about the presented real world entity. Within a module, whatever representations have been produced are meant to be veridical and hence can be simply taken as true by subsequent processing within the module. The situation is somewhat different for central processing, because central processing is by assumption, non-encapsulated and hence inferentially promiscuous. While the modules would have evolved to deliver generally true representations to the central processor, there will be occasional misrepresentations. If a system then produced an indefinite number of inferences from its existing beliefs ad lib, there would be chaos in the ensuing mess of true and false beliefs. The system would need some form of house-keeping, some way of checking for contradictions, and then defeasibly labelling some beliefs as appropriate for acting on and other beliefs not. That is, the system requires its representations to have their factuality explicitly marked. Factuality explicitness would be natural for inter-modular communication but not for intra-modular communication. The perhaps largely predication-implicit representations within modules and the largely fact-explicit representations between modules allow two corresponding different types of control, which we call vehicle-control and content-control, respectively (Dienes and Perner 1999, p. 744; Perner in press). In vehicle-control the explicit representation of mere properties is sufficient to trigger (according to established routines) the next appropriate operation. Absent-mindedly driving to work would be an example of such featural triggering we call vehicle-control. In this case, there is no representation of the required input-output mapping that exists separately from the processor that performs the mapping. But adult humans are able to do more. We can sometimes override a normal routine, by specifying an imperative, "Turn right!", that contradicts what we would normally do in the circumstances, or we can be controlled by a representation of a conditional rule that controls behaviour with no previous practice "If they say the plane is cancelled, then turn back." These representations must have fact-explicit content because they describe not what should be taken as a fact, but what should be done. And they can come to control just those processors that implement the content specified. Fact explicit representations are necessary for this flexible control we call content control. Thus, in general, it will only be representations formed in the central workspace that can come to exert content control; intra-modular control will largely be vehicle-control. If information exchange across modules does largely involve explicit factuality an interesting question arises whether these processes are also conscious. We have pointed out that explicit factuality is very close to conscious awareness (Dienes and Perner 1999, p. 741); under normal conditions in adult humans, occurrent fact explicit representations become fully attitude explicit, e.g., "I know that this is a cat," satisfying the necessary and sufficient conditions of conscious awareness according to some philosophers (Carruthers 2000; Rosenthal 1986). If Frawley's claim that across module linguistic processing requires explicit factuality can be substantiated then this might shed some interesting light on these issues about the basis of conscious awareness. Intuitively one would think that we are typically not consciously aware of the information interchange between the phonological and the morphological module, indicating the use of fact implicit representations. We should be consciously aware of exactly those kinds of linguistic features that are based on an exchange between modules using fact explicit representations. Frawley can be read to think that we are conscious of these features: "To put it another way, if you have no access to the occasioning of a piece of information, you may have the implicit feeling that it did happen, but you do not know that it did" (p. 47). Knowing that the across module processes happened suggests conscious knowing. But that, presumably, still awaits systematic investigation.

REFERENCES

BAARS, B. J. 1988. A Cognitive Theory of Consciousness. Cambridge University Press, Cambridge, UK.

CARRUTHERS, P. 2000. Phenomenal Consciousness: A Naturalistic Theory. Cambridge University Press, Cambridge, UK.

DIENES, Z., and J. PERNER. 1999. A theory of implicit and explicit knowledge (target article). Behavioral and Brain Sciences, **22**:735–755.

FODOR, J. 1983. The modularity of mind. Bradord Books, Cambridge, UK.

FOLEY, M. A., M. K. JOHNSON, and C. L. RAYE. 1983. Age-related changes in confusion between memories for thoughts and memories for speech. Child Development, **54**(1):51–60.

KOWALSKI, R. 1979. Alogrithm = logic + control. Communications of the ACM, **22**:424–438.

PERNER, J. In press. Dual control and the causal theory of action: The case of nonintentional action. *In* Agency and Self-awareness. *Edited by* N. Eilan and J. Roessler. Oxford University Press, Oxford, UK.

ROSENTHAL, D. M. 1986. Two concepts of consciousness. Philosophical Studies, **49**:329–359.