The metacognitive implications of the implicit-explicit distinction

Zoltan Dienes & Josef Perner
University of Sussex and University of Salzburg

Abstract In this chapter we establish what it is for something to be implicit. The approach to implicit knowledge is taken from Dienes and Perner (1999) and Perner and Dienes (1999), which relates the implicit-explicit distinction to knowledge representations. To be clear about exactly what our claims are we first discuss what a representation is, what it is for a representation to represent something implicitly or explicitly and apply those concepts to knowledge. Next we show how maximally explicit knowledge is naturally associated with consciousness (according to the higher order thought theory). Then we discuss the relationships between explicit knowledge and metacognition, where metacognition is considered in terms of both its monitoring and control aspects, to shed light on conscious and unconscious perception, episodic memory, and volitional control. We will then show how implicit learning should be viewed in metacognitive terms, and conclude that people's relative lack of metaknowledge in implicit learning paradigms justifies the claim that people have acquired genuinely implicit knowledge.

In Izaute, M., Chambres, P., Marescaux, P.-J. (Eds), Metacognition: Process, function, and use. Dordrecht, Netherlands: Kluwer Academic Publishers, forthcoming.

1. INTRODUCTION

In this chapter we will consider the relation between the implicit-explicit distinction and metacognition (Reder, 1996). To understand this relationship we will need to first consider what a representation is, because we subscribe to a representational theory of knowledge; i.e. we consider that when a person occurrently knows something, that is because they have formed a representation (be it connectionist or symbolic or something else) about what they know. We indicate how a representation can represent different contents implicitly or explicitly, and use this to derive a hierarchy of explicitness of knowledge (Dienes & Perner, 1999; Perner & Dienes, 1999). Explicitness will then be related to consciousness via the higher order thought theory. With this framework in place, we can finally consider metacognition in its monitoring and control aspects, and then look at the metacognitive approach to implicit learning.

2. REPRESENTATIONS AND CONSCIOUSNESS

In order to clarify the relation between metacognition and the implicit-explicit distinction, we will first need to be clear about what representations are, and how they are related to consciousness. One can find considerable disagreement in the literature about the relation between representations and consciousness. For example, on the one hand, Whittlesea and Dorkin (1997) asserted that people in general "do not have direct, conscious access to those representations" that underlie performance on tasks (p. 64); on the other hand, Dulany (1996) and Perruchet, Vinter, and Gallego (1997) believed that all mental representations are

conscious. In order to make meaningful claims of either sort, we must first be clear what we mean by representation.

So what is a representation? Consider an unambiguous case of a representation: A map of a town. In this case, and in general, a representation consists of something physical (the representational medium, for example, paper and ink) that is about something else (the representational content, for example, the town). But how is it that an object - paper and ink - can acquire meaning, a content? Or consider a case closer to psychology. How could, say, a pattern of firing of a group of neurons in a person represent a cat? You might suggest - taking note of the way that neurophysiologists determine what a cell, or group of cells, code - that the pattern represents a cat if it is correlated with the presence of cats: Whenever you show a cat to the person, those neurons fire. Unfortunately, this does not quite do; it does not allow the person to misrepresent. If he saw a skunk on a dark night, the same neurons might fire. On the correlation story he has not misrepresented the skunk as a cat; he has just correctly detected a cat-OR-skunk-on-a-dark-night. But representations can misrepresent and any theory of representation must allow for that.

Correlations between patterns of neural activity and cats arise in people due to an evolutionary or learning history that has selected that pattern of activity because of the function it performs in dealing with cats. One might say the pattern has the function of indicating cats; or the function of producing further internal or behavioural reactions appropriate for dealing with real or imagined cats. According to one dominant (and we think persuasive) approach in philosophy, representations represent something precisely because of the functional role they play. Thus, on Dretske's (1988) approach, if A has the function of indicating B then A represents B. For example, if a pattern of neuronal activity has the function of indicating cats, then it represents "cat". If it fires because of a skunk on a dark night, then it has misrepresented the skunk as a cat. Function can be produced by evolution or learning, or, in the case of artifacts like a map, by our intentions.

Is there any reason why all representations, thus defined, should be conscious? Not at all, maps are not conscious. Imagine building a robot to interact with the world, and the robot will be conscious of some aspects of his world. It may be useful to have the activity in some circuit have the function of indicating a particular internal or external state of affairs. There seems to be no a priori reason why the content of all such representations should constitute the content of the robot's conscious experience. Perhaps the representation was useful simply temporarily to inform another process downstream of processing; or the problems it is used to solve are "local" problems that do not need to concern the processing system generally. In any case, the extent to which people have interesting unconscious representations is an open question, and an empirical question given a theory of consciousness.

The relationship between consciousness and representation may be partly open but the relationship is not one of complete independence. Our conscious states are typically characterized by being about things (Brentano, 1874); thoughts are always about what is thought, desires are always about what is desired. Some argue that all conscious states are about something (e.g. Tye, 1995), but it is enough to note that many conscious states are about something. Given a materialist theory of the mind, it follows that conscious states must be (in at least many cases) representational, because the states have a physical embodiment (the representational medium; that is, part of the brain) and are about something else (the representational content, the content of the conscious state). So the content of consciousness is just the content of some representation. In this sense we can say that at least some mental representations are conscious.

What makes some representations conscious (when others are not)? One might answer that all mental representations are conscious (e.g. Perruchet et al, 1997). One then needs an account of what makes a representation "mental". "Mental" has been defined as any state that could in principle become conscious (Searle, 1990). For example, states of detectors in the liver signalling the presence of glucose could not become conscious states, so they are not mental. Their unconscious status is no more mysterious than a map not being conscious

(though the consciousness of mental states would remain mysterious). How does one view unconscious perception or unconscious learning if one assume that all mental states are conscious? A possible argument is that subliminal perception is not possible, because perception is a mental state, and all mental states are conscious. We believe this position has been falsified by the evidence (e.g. Debner & Jacoby, 1994; Merikle & Joordens, 1997). But a retreat position is available: the perceptual states controlling behaviour in a subliminal perception experiment are obviously ones that can not be made conscious, so they are not mental. Since this retreat position is always available (unless further stipulations about what counts as mental are made) whether the experimental evidence supports subliminal perception or not, it is not clear to us what work is being done by the claim "all mental representations are conscious". We will argue that some representations controlling behaviour are conscious, and that it is possible and testable that other representations controlling behaviour (occurrent representations resulting from perception, occurrent representations resulting from learning) are unconscious. In order to make this argument, we will need to distinguish implicit from explicit representations (according to the framework of Dienes and Perner, 1999) and employ a theory of consciousness (the higher order thought theory). Finally, we will be in a position to discuss how metacognition (the monitoring and control of such representations) is related to the implicit-explicit distinction.

3. IMPLICIT VERSUS EXPLICIT REPRESENTATION AND KNOWLEDGE

According to Dretske (1988), if it is the function of state A in a representational medium to indicate B then A represents B. A has the function of indicating B partly because the state of A is used as information by the rest of the system to respond appropriately to B. Now for A to indicate anything, for it to be used as information, requires that at a minimum that the representational medium can go into two states. For example, if A represents “cat”, then there should be one state for “cat” and another state for “not a cat” or “uncertain if cat or not-cat”. We will define the explicit content of a representation in the following way: Distinctions (e.g. cat/not-cat) are explicitly represented only if there are corresponding distinctions in the representational medium. However, the explicit content of a representation rarely constitutes its entire content, as we will now begin to see.

A representation may express content that has a structure. But there is no reason why all the elements and relations in that structure must themselves be explicitly represented. Consider a device for distinguishing different animals. If you put a cat in front of its sensors, it goes into a “cat” state; if you put a dog there, it goes into a “dog” state, and so on. Thus, the distinction between cat and dog is explicitly represented, because differences in the device’s representational medium correspond to the different animals placed before it. But the full content expressed by the representation when the device goes into its “cat” state is more than just “cat”; rather the device is indicating (and has the function to indicate) at least that “this is a cat”. We could not say anything less, for example, that it only expresses knowledge of cat-ness, or of the concept of cat. The device can convey information that “this is a cat”, or “this is a dog” by going into different states. Yet, what are made explicit within the vocabulary of this device are only the properties of being-a-cat, being-a-dog, etc. That it is “this” rather than “that” object that is a cat is an element of the structure of the expressed content, an element that helps constitute the meaning of the representation, but there is no difference in the representational medium that corresponds to “this” rather than “that”.

Based on the foregoing logic, we will distinguish explicit representation from something that is only implicitly represented in the following way: Any environmental feature or state of affairs that is not explicitly represented but forms part of the representational content is represented implicitly. Thus, in the example of the animal detector, the animal is represented

explicitly, but the fact that it was this animal is represented only implicitly. To give another example, the function of a bee dance is to indicate the location of nectar; this is its representational content. It represents the direction of the nectar explicitly, because the angle of the dance varies systematically with the direction of the nectar. However, the fact that it is about nectar (see Millikan, 1993, for the argument that the bee dance is indeed about nectar) is represented only implicitly. We will now apply the implicit-explicit distinction to what it is to have knowledge.

What is it to have knowledge? First there is the content of the knowledge: A proposition, i.e. something that can be true or false. This usually involves predicating a property (e.g. "is bald") to an individual (e.g. "the king of France").[1] Second, the content must be a fact at a given time. Third, there is a person ("I") having an appropriate relationship to this proposition, i.e. a relationship of knowing rather than, for example, wishing, guessing, considering or dreaming.

A representation functioning as knowledge need not make all this structure explicit. The following does constitute a fully explicit representation of the knowledge that the present king of France is bald "I know (that it is a fact) that the present king of France is bald". We will now consider ways in which a person may not represent this state of affairs fully explicitly, according to the taxonomy described by Dienes and Perner (1999).

At one extreme, the person may explicitly represent only a property of a presented object or event. For example, when a person is flashed the word "butter", during perception of the event they may not form an explicit representation of the full proposition "The word in front of me has the meaning butter". Instead the meaning butter is activated but it is not predicated of any particular individual (i.e. "the word in front of me"). The representational medium contains no distinction that indicates different individuals. So the full content of the proposition is not made explicit. But if the person reliably acts appropriately towards the stimulus (in a certain context) the representation is functioning as knowledge. Thus, its status as knowledge, the fact that the feature applies to a particular individual (presented word) is implicitly represented, by our definition. This is maximally implicit knowledge on our scheme. Consider for example a blindsight patient presented with a square or a circle in their blind field. They can reliably indicate whether the object is a square or a circle, but provide no evidence that anything more than "square" or "circle" has been explicitly represented about the fact that it is a square or circle presented to them (e.g. Weiskrantz, 1988).

We suggest that under subliminal conditions only the properties of a stimulus (the kind of stimulus) get explicitly represented (e.g., the word "butter"), not the fact that there is a particular stimulus event that is of that kind. This would be enough to influence indirect tests, in which no reference is made to the stimulus event (e.g., naming milk products), by raising the likelihood of responding with the subliminally presented stimulus ("butter" is listed as a milk product more often than without subliminal presentation). The stimulus word is not given as response to a direct test (e.g., Which word did I just flash?) because there is no representation of any word having been flashed. Performance on a direct test can be improved with instructions to guess, because this gives leave to treat the direct test like an indirect test, just saying what comes to mind first.[2]

---

[1] This is true, even of procedural knowledge. A procedure is of the general form "If condition X, then action Y". In a calculator, it may be: If "5 X 6" then show "30". The property of being 30 is predicated of the result of the operation 5 X 6. Note also that detailed perceptual properties can be predicated of individuals.

[2] It is the fact that the person can reliably identify the actually presented word (when e.g. given leave to guess) that entitles us to say the person has knowledge, and therefore allows us to talk about implicit knowledge. It is only in an appropriate supporting context that the representation functions as knowledge of a particular event. Nonetheless, we will loosely refer to the representation as providing implicit knowledge in all contexts. In many cases (e.g. Bridgeman, 1991; see Dienes & Perner, 1999), the visual system evolved the use of such (implicit) representations precisely because of their role in such supporting contexts, and so the proper function of the representation is indeed knowledge.

At the next stage of explicitness, the person represents the full content of the proposition (i.e. including the individual that the property is predicated to) and then represents the temporal context of the fact and whether indeed it is a fact or not. This extra representation of time and factuality may seem gratuitous, but it is important for explicit memory rather than mere implicit memory (which can be just be based on maximally implicit knowledge, where just a property is represented explicitly): To recollect the past one must represent the past events as having taken place in the past.

At the final stage of explicitness, one represents that one knows a particular proposition. For example, in the case of perception, the knowledge is based on seeing and the perceptual process may yield the representation "I see that (it is a fact that) the word in front of me is butter". This representation would enable a person to confidently report on seeing the word butter; in other words it would enable conscious perception, as we will now see.

4. HIGHER-ORDER THOUGHT THEORY OF CONSCIOUSNESS

What would make the perception of the word in front of you being butter a conscious perception? In general, under what conditions is a mental state (sensation, thought, desire, etc) a conscious mental state? We will answer this question by reference to the higher order thought theory of consciousness (e.g. Armstrong, 1980; Rosenthal, 1986, 2000a,b,c; Carruthers, 1992, in press), in particular Rosenthal's higher order thought theory, a philosophical theory of consciousness we find appealing for its simplicity and elegance. In order to have an account of a mental state, like a thought, being conscious, we need to consider the logical possibility of thoughts being unconscious, so we can consider what would make the mental state conscious independently of simply being a mental state. That is, to say that someone is thinking, we should not presume that they must be consciously thinking; they could be unconsciously thinking. With that proviso in mind, we can consider how we become conscious of events and things. In general, I can be conscious of things in two ways; by perception and by thinking. I can be conscious of a problem by thinking about a problem; I can be conscious of you by seeing you or just by thinking about you being there. If we flash a person either the word "butter" or the word "grass", and they can later make a forced choice discrimination above chance about the identity of the word, we can say he is conscious of the word because he saw the word. But by "conscious of the word" we do not necessarily mean consciously aware of the word or that he beheld the word with a conscious mental state. In a sense he is conscious of the word; but the seeing itself need not be a conscious mental state. For a mental state to be a conscious mental state, we should be conscious of the mental state. We could not claim that a person has a conscious mental state, and also claim that the person is not conscious of being in the mental state. According to Rosenthal, the relevant way of being conscious of the mental state is to have a thought about the mental state. For example, if the mental state is seeing that the word is butter, one becomes conscious of the mental state by thinking "I see that the word is butter"; because the state of affairs of the word being butter is now beheld with a conscious mental state, the person is consciously aware of the word being butter.

In general, according to these theories, it is a necessary and sufficient condition for conscious awareness of a fact that I entertain a second order thought that represents the first order mental state (in the example, the first order mental state is seeing that the word is butter; the second order thought is representing that I am seeing that the word is butter). But this is just the same as our requirement for knowledge to be fully explicit: The person must represent that they know (for example by seeing) that the word is butter. Our framework shows why explicitness is often intuitively felt to have something to do with consciousness.

The second order thought does not make itself, the second order thought, conscious, it just makes the first order thought that it is about conscious. The second-order thought "I see that

the word is butter" only makes one consciously aware of the word being butter, not the fact that one sees that the word is butter. To be aware that one knows it by seeing, there needs to be a third-order thought that makes the second order thought conscious. Typically, when we consciously know a fact, we also know how we consciously know it. Normally, one could not sincerely claim "I am conscious of the word being butter" and at the same time deny having any knowledge of whether one sees the word, or hears about it, and so on. This fact provides strong evidence for higher order thought theories in humans. Presumably whatever mechanism produces second order thoughts is just the same that produces third order thoughts; and it would seem highly plausible that if a representation was available to the mechanism for second order thought, the output of the mechanism would be available for third order thoughts. At least, that would seem to be the simplest way for evolution to have set things up[3]. Thus, typically, according to the theory and as supported by the facts, one would expect people to be able to say how they are currently aware of something if they are able to say they are currently aware of it at all[4].

5. METACOGNITION: MONITORING

According to Nelson and Narens, (1990) metacognition has both monitoring and control aspects to it. In this section we will consider metacognition as monitoring; in the next section we will consider the control aspect of metacognition. Metacognition literally means cognition about cognition. There is thus an obvious relation to higher order thought theory (a link discussed by Rosenthal, 2000a,b,c), since the latter claims that conscious mental states arise exactly from thinking about thinking. Fully explicit knowledge, in our sense, is thus a case of metacognition; what implicit representations lack is metacognitions about them (cf Kinoshita, this volume). Paris (this volume) discusses when metacognitions are harmful, benign, or useful; one use, of which we can be grateful, is in making us consciously aware[5]. Every moment of our waking life we are engaged in automatic and unconscious metacognitions providing us with all our conscious experiences. This is the pervasive sea of metacognitive monitoring in which we live.

Rosenthal argues that higher order thoughts make us consciously aware when they are (a) assertoric (they authoritatively assert that we are in a mental state); and (b) appear unmediated; that is, they are not the result of a conscious inference. If in a subliminal perception experiment, a subject thinks "I did not see anything. But, since I am forced to guess, 'butter' comes to mind easily. I must have seen the word butter", there is a higher order thought about seeing the word butter. But the thought arises as a conscious inference, and thus it does not make the subject conscious of seeing that the word was butter. The subject may just be conscious of inferring or guessing that the word was butter, but not conscious of seeing. An unmediated thought to the effect that they are guessing is formed, but not to the effect that they are seeing.

Further, according to our arguments, if a subject in a subliminal perception experiment when flashed the word "butter" just forms the maximally implicit representation "butter", the

---

[3] According to Carruthers' (1992) potentialist higher-order thought theory, a representation is conscious if it is recursively available for successively highly order thoughts. It directly follows from this account that if you are conscious of X you are also potentially conscious of how you know X.

[4] This observation shows that the higher order thought theory is not just a conceptual analysis of how words are used; it has genuine explanatory power. Another illustration of the explanatory power of the higher order thought theory is that it corresponds to the measure of consciousness Cheesman and Merikle (1984; 1986) called the subjective rather than the objective threshold; it is the subjective threshold that appears to divide qualitatively different psychological processes.

[5] This begs the question of why second order thoughts may be useful from an evolutionary perspective; Miller (2000) argues eloquently that maybe natural selection had little to do with it; it may have been predominantly sexual selection.

subject does not have any conscious experience caused by the representation. According to Dulany (1996) and Tzelgov, Ganor, and Yehene (1999), the formation of any semantic representation (e.g. the maximally implicit representation "butter") is sufficient for conscious experience. If, for example, one forced the subject to choose a word that may have been just flashed the subject may choose "butter" at above chance rates; therefore, the argument goes, the subject must have consciously seen the word butter. In this situation, according to us, the mechanisms brought into play by the attempt to guess use the unconscious representation "butter" to make a guess, and thus make it conscious as a guess. But the subject is not conscious of the seeing as seeing, so in this sense we can say the subject saw the word only unconsciously. (Further, conscious awareness of the word even just as a guess was only brought about by probing for the word; it was not automatically produced by the act of seeing.)

Searle (1983) argued that when we see an event, we experience the event as directly causing the visual experience. We don't have to follow a chain of reasoning to know the event caused the experience; the knowledge that the event caused the experience is part of the experience itself. That is, you know non-inferentially that the word "butter" directly caused your knowledge that the word butter was there. This is consistent with our claim that inferentially guessing cannot be regarded as a case of conscious seeing. But if the non-inferential understanding that the event directly caused the visual experience was necessary to all conscious seeing (as opposed to e.g. guessing what you must have seen), then it seems to follow that young children and animals do not consciously see (an implication pointed out by Armstrong, 1991), because they cannot make the conceptual distinction between experience and reality needed to understand that the experience was caused by the reality (Flavell, Flavell, & Green, 1983). The answer is that conscious seeing does not require this understanding; it just requires one to think "I see that the word is butter" with conviction and in a way that appears unmediated. Given an adult's understanding of seeing, "seeing that the word is butter" will mean to the adult that that the relevant state of affairs in the world - the word on the screen being butter - caused the visual experience, and this fact itself will appear (at least on reflection) to be part of the visual experience (giving vision what Searle calls a "self-referential" nature). That is, self-referentiality is part of how the adult understands vision; so when the adult thinks "I see that the word is butter", the self-referentiality is implicit in the use of the representation "see". However, the self-referentiality does not need to be explicitly represented in each episode of conscious seeing. It will be explicitly represented whenever the adult reflects on the seeing process; thus, it will seem to the adult that the experience of self-referentiality is always part of seeing. In fact, it only arises when the adult forms appropriate third-order thoughts. It is only there when the adult looks for it; thus, it appears to be always there. The child or the animal do not need to understand that seeing works this way; they just need some more primitive concept of seeing. Thus, children and animals, by using such concept of seeing as they do have, can have conscious visual experiences (but only when they use this concept in an assertoric way that does not appear mediated to them).

If one merely thought "I see that the word is butter", one would consciously see that the word is butter, but only in a conceptual way. Normally visual experiences have content that cannot be exhaustively described by concepts (e.g. Chrisley, 1996); for example, in looking at an apple, one may experience that the apple has a fine-grained shade of red for which one has no concept. This content not captured by the concepts the person actually possess has been called non-conceptual content (Cussins, 1992). Visual experiences normally have distinctively visual non-conceptual content. According to Rosenthal, we are conscious of our experiences only in the way they are represented to us by our higher order thoughts. Thus, when we are consciously aware of non-conceptual content, we must have predicated the non-conceptual content to the relevant object or event we are beholding and formed a higher order thought to the effect that we are seeing that non-conceptual content. The non-conceptual content becomes part of the higher order thought; only in this way could we be consciously

aware of the non-conceptual content. Because such visual experiences have distinctively visual non-conceptual content, reflection on those experiences leads one to think that they are obviously visual experiences. But once again, one is consciously aware of the visual nature of those experiences only when one reflects on them with relevant third-order thoughts. It will seem to us that their visual nature is always apparent to us, because whenever we check, it is apparent to us for reasons (the presence of relevant non-conceptual content) that have been true all along.

Everything we have said above about perception applies to memory, with the necessary changes. For us to have an episodic memory of seeing butter on the list we must think the second-order thought[6] "I remember that I experienced that butter was on the list" with conviction and in a direct and unmediated way; i.e. the truth of the thought does not appear to the person to be known as a consequence of other thoughts and events. Such an authoritative unmediated thought is sufficient for us to know consciously something happened as part of our personal past. Genuine episodic memory also involves us being aware of the act of memory as an act of memory, and this involves forming a relevant third order thought to the effect that one knows one is remembering; just as in the visual case, where a third order thought is necessary to be aware one is seeing. Just as in vision where most acts of conscious seeing may only involve relevant second order thoughts, many acts of episodic memory (particularly when one is engrossed in memory) may involve only second order thoughts; but the third order thought will be generated whenever internally or externally probed for and help constitute the full experience of remembering.

Dokic (1997; see also Perner, 2000a) considered a case where a person believes they have experienced an event, but wonders if he believes this because he is really remembering or because he was told as a child. He asks his parents, and the parents assure him that he really experienced the event and no-one could have told him. So the person forms the representation "I remember I experienced the event", but this does not seem to be a genuine case of episodic memory. The reason why it does not seem to be genuine is for the same reasons guessing in subliminal perception experiments is not genuine conscious seeing. In the memory case, the person initially believes the thought "I experienced the event" may have been known by being told, and thus it is not a case of remembering one experienced the event at all. The parents' later comment leads the person to think "I remember I directly experienced the event" only as a conclusion derived from other people's comments, and one is conscious of its inferential nature (as in our vision example). If the person later forgot the conversation he had with the parents, but now experienced the thought "I remember I experienced the event" in what seems to him to be a direct and authoritative way then he would experience knowledge of the event as an episodic memory. If in fact the event never happened (he and his parents were wrong that it had happened) then he would still have an episodic memory, albeit a false one.

Adults understand memory in a self-referential way (Searle, 1983), they understand remembering must involve a real event directly causing the memory, and that the knowledge that the real event caused the memory must itself be caused by the event. This understanding of remembering can be implicit in the meaning of "remember" and need not be explicitly represented on every occasion something is remembered; the conscious awareness of having experienced a past event comes simply from an assertoric and non-inferential thought to the effect that one is remembering. A further relevant assertoric and non-inferential third order thought that one knows one is remembering provides the conscious awareness that one is remembering. As further argued by Perner (2000a), genuine recollections will involve representing the sensory content (and hence the non-conceptual content) involved in experiencing the event, representing it as part of the remembered event. This later step (combined with a relevant third order thought) creates what would be a conscious memory on

[6] This may seem like a third order thought because one thinks that one remembers that one experienced; but the "experienced" is past tense and so not an occurrent mental state, but simply a fact of the past.

Jacoby's (1991) account; and a "remember" rather than a "know" response according to the remember/know procedure (Gardiner, 1988; Tulving, 1985).

6. METACOGNITION: CONTROL

In this section we discuss metacognition as a control process. As well as providing us with all our conscious experiences, another useful aspect of metacognition is that it also enables all acts of volitional control, as we shall now see. Almost continuously throughout our waking life we are engaged in automatic and unconscious metacognitions providing us with volitional control over our actions and mental processes. This is the pervasive sea of metacognitive control in which we live.

When we cognitively control our own cognitions, we are engaged in metacognition. Voluntary control is an example of metacognition. Voluntary or intentional control of knowledge means that one can use it intentionally. That is, one needs to represent that one intends to use that knowledge. One needs to reference the appropriate response as *something intended* and not, for example, as an existing fact. Thus, the factuality (or otherwise) of the content of the knowledge and the mental state by which one considers the content (i.e. desire) must be made explicit. In performing a voluntary action, the action is voluntary by virtue of forming a higher order thought to the effect that one is intending the action; implicit in the meaning of intending for adults will be the understanding that the action is performed by way of carrying out the intention of performing that very action (cf Searle, 1983). This analysis shows why the common notion that voluntary control is associated with explicitness is justified. Voluntary control is also shown to be essentially a metacognitive process, to involve second order thoughts, and hence consciousness[7].

Perner (1998, in press b) presented a dual control model of action, in which there are two levels of control, vehicle and content control. Control of action can occur just at the level of representational vehicle: An action schema comes to control behaviour simply because of the existing associative links between the representation of current actual conditions - the schema's triggering conditions - and the production of the action. In this case, the action schema that controls behaviour is the one with most activation, and here activation is a property of the representational vehicle; the degree of activation does not represent the content of the schema, it just determines the probability with which it will control behaviour. In contrast to vehicle control, control of action can occur at the level of representational content. A representation is formed (e.g. from verbal instructions or mental planning) of the required mapping between conditions and actions ("if condition C then do action A") or simply of the desired action ("do A"). In content control, it is the content of this representation that determines which schema comes to control behaviour; that is, the schema with the conditions and action described by this representation. This representation must represent conditions and actions at least fact explicitly because it states a hypothetical "If condition C…"; it should not in itself lead to registering that condition C has obtained, it simply states what to do if condition C were to obtain[8]. The representation also represents the action-to-be-

[7] There is an interesting symmetry with perception: The action must seem to be caused by the intention (higher order thought) in a way that appears unmediated; if the action appeared mediated, it would not be a voluntary action but the outcome of a voluntary action. Voluntary action *requires* unconscious processes of mediation, just as conscious perception requires the mechanisms mediating between first order mental states and higher order thoughts to be unconscious.

[8] Note that procedural knowledge - often represented in the form of procedures like "If condition C, then do action A" (e.g. Anderson, 1983) - does not require fact explicit representation. In fact, declarative knowledge differs from procedural knowledge precisely because declarative knowledge declares what is the case, i.e. represents factuality explicitly, whereas procedural knowledge need not. "If condition C, do action X" is a declarative representation of what a procedure may represent fact implicitly by virtue of implementing the right links between conditions and actions. Whenever, for example, the condition is occurrently represented in the

performed as something needed and is therefore not actually a fact; the goal state of the completed action must therefore be represented fact explicitly. In content control, the representation of the appropriate condition-action mapping causes the relevant action schema to control behaviour, regardless of the existing strength of associative links between current actual conditions and particular actions. In vehicle control, there are the conditions and actions represented in the schema, which do not need to be represented fact explicitly; in content control, there is, in addition to the representations embodied in the schema themselves, the occurrent fact-explicit representation of the required actions that determines schema choice. For example, consider driving from work to the supermarket and the route taken is in part the same as the more normal route from work to back home. If one did not keep actively in mind the new action required at a crucial juncture (so content control fails) one would end up driving home (vehicle control determines behaviour; the action most strongly associated with current conditions is performed). The implicitly acquired control described by Reder (this volume) is an example of vehicle control.

Some tasks (the executive function tasks described by Norman and Shallice, 1986) necessarily involve content control, for example, inhibiting normal reactions in order to do something novel in a situation (Perner, 1998). This type of task requires one represent the novel action as something required, and therefore it must be represented fact explicitly. Conscious intentions use content control, because they represent desired condition-action mappings, i.e. they use fact-explicit representations to control schema choice. Conversely, vehicle control does not require conscious intentions. For example, Debner and Jacoby (1994) flashed a word to subjects and then asked them to complete a word stem with anything EXCEPT the word that had been flashed. The conscious intention to not use that word could inhibit the action schema responsible for completing with that word and allow other action schemata to control behaviour. Thus, for words flashed for a long enough duration, stems were completed with those words at below baseline levels. However, if words were flashed very quickly, they were not consciously perceived, no conscious intention could be formed that inhibited their normal use, and an action schema was chosen simply based on which became activated most strongly by the triggering stem. That is, only vehicle control was possible. In this situation, subjects completed stems with the flashed words at above baseline levels. (Of course, control occurs in the context of a hierarchy of goals; even vehicle control is relative to this context. Subjects would have had content control of the general action "complete the stem with **some** word".)

Content control only actually requires fact explicit representation; it does not require full explicitness, so it does not actually require conscious representations of required condition-action mappings. Perner (2000b) pointed out that our framework predicted the possibility of content control (i.e. the control required in executive function tasks) without conscious awareness. This seemed an unlikely prediction and led Perner to suggest the framework should be sent back to the drawing table. In fact, however, the predicted existence of content control without awareness is confirmed by hypnosis and related psychopathological states like hysteria, and everyday dissociative phenomena, which therefore provide supporting evidence for the validity of the framework. Sheehan and McConkey (1982) and Spanos (e.g. 1986) have emphasized the strategic goal-directed nature of hypnotic responding. A subject can be given a suggestion to count but always miss out the number "4". The inhibition of normal associations are required, so content control is required. Nonetheless, susceptible subjects will respond successfully to the suggestion (counting "1,2,3,5,6,…"), all the while affirming their ignorance that they are doing anything strange. Similarly, Spanos, Radtke, and Dubreuil (1982; Spanos, 1986) found that highly susceptible subjects suggested to forget certain words in any type of task given to them produced those words at a below baseline level in a word association test. This performance again calls for

---

procedure, the presence of the condition would be treated as a given because its factuality is taken for granted and the procedure would apply the action.

content control because the existing associations that would be produced by vehicle control must be suppressed. In general, virtually any arbitrary behaviour can be hypnotically suggested despite the fact that such behaviour might be novel to the person; it is highly plausible that many hypnotic responses are under content control. Yet highly susceptible subjects claim that their actions do not feel like normal consciously controlled actions; they seem strangely involuntary. And indeed they would seem involuntary if one had not represented the relevant goals as things to which the "I" had a mental-state relation (Kihlstrom, 1997), i.e. if ascent from explicit representation of factivity to full explicitness had been inhibited[9].

Content control of actions might be easier if one kept in mind not just declarative representations of the content of goals and condition-action mappings, but also representations of the appropriate mental states "I wish that…". That is, content control might be easier if performed with awareness rather than without. Perhaps the extra representations of mental states and the use of the "I" representation allows extra activation sources to feed to the controlling structures and support the controlling fact-explicit representation (cf Anderson 1983; Kihlstrom & Cantor, 1984). A person particularly skilled at content control may be most able to engage in it even when the I and mental states are not being represented. That is, such a person may be particularly able to experience hypnotic effects. In sum, the prediction is highly hypnotizable subjects should be better than low hypnotizables at tasks requiring content control. Indeed, there is a large body of evidence for this claim; for example, highs can generate random numbers with a greater degree of randomness than lows (Graham & Evans, 1977; this is regarded as an executive task, Baddeley, 1986); and in selective attention tasks highs can select on the basis of representational content (semantic selection, or "pigeon holing", Broadbent, 1971) to a greater degree than lows, but they cannot filter according to purely sensory features any better than lows (Dienes, 1987).

7. IMPLICIT LEARNING

The term implicit learning was coined by Reber (1967) to refer to the way people could learn structure in a domain without being able to say what they had learnt. Later, Broadbent and Aston (1978) independently applied the term "implicit" to such knowledge. Reber had looked at the way people learned artificial grammars; Broadbent looked at the way people learned to control dynamic systems. Both Reber and Broadbent found that people could make appropriate decisions (in deciding on grammaticality and setting the value of a control variable, respectively) without being able to explain why their decisions were correct; and both intuitively felt that the word "implicit" captured the nature of this learning. But what is implicit about implicit learning?

---

[9] The absence of second order thoughts would preclude the formation of a second order thought such as "I perform this action of raising my arm by way of carrying out this intention (of mine)". It is the absence of this "intention-in-action" (Searle, 1983) that makes the act feel involuntary. This analysis of hypnotic responding - appropriate first order control states in the absence of corresponding second-order thoughts - may account for many suggestions, like motor suggestions and some strategically-mediated cognitive ones. In addition, hypnotic hallucinations appear to rely on a complementary state of affairs: second order thoughts in the absence of the corresponding first order perceptual states. One may experience being in a mental state (seeing, feeling pain) even though one is not really in it (as discussed by Rosenthal, 2000a,b,c for some non-hypnotic contexts). The prediction is that such illusory second-order thoughts should arise most often when: there is strong expectation that one will have the first order and higher order states; the person has vivid imagery and capacity for imaginative involvement; and the subject engages in appropriate fantasy simulations of the first order state (to produce sufficient first order information to trigger the primed second order thoughts). These are indeed important predictors of hypnotic responding (e.g. Kirsch, 1991; Spanos, 1986). A third route to hypnotic responding may be in the creation of different "I"s (Kihlstrom, 1997), but the complexity of this route would presumably restrict its use to very few people.

If people could describe the knowledge they had acquired the knowledge would have been represented at least fact explicitly. Anything we can state verbally we can consider whether it is true or not; hence all verbalizable knowledge is at least fact explicit. Further, by expressing the knowledge verbally, a person can consider their relation to the knowledge; if they correctly know that they know it, then the knowledge is fully explicit according to our framework. In order to test a hypothesis, a fact explicit representation must be considered ("If X is true, then…"), because conditional and counterfactual statements necessarily involve explicit consideration of factuality. Seeing why a hypothesis passes whatever test is set, is to see why the validated hypothesis is now part of one's knowledge. Hypothesis testing (when considered as such by the system that does it) is explicit learning. In contrast, the knowledge produced by implicit learning has not been represented as knowledge by the learning process. It's status as knowledge is left implicit in its functional role. The knowledge was acquired by the system in order to facilitate the very task the subject is engaged in; this is what makes it knowledge. For example, in the dynamic control tasks of Berry and Broadbent (1984), people appear to learn what actions lead to the goal in different specific situations to form a look-up table; and this look-up table determines future actions in the same situations (Dienes & Fahey, 1985). People do not explicitly remember these situations (Dienes & Fahey, 1998), they just respond appropriately to them.[10] People can respond appropriately without knowing they have knowledge; this lack of metacognition is what makes the knowledge implicit. Further, the knowledge need not be represented as factual or not; it's factuality is left implicit in the way it is simply taken as true. That is, according to our framework, implicit learning does indeed produce implicit knowledge. At least, this is what we believe the experimental evidence has shown, as we now describe. We will consider artificial grammar learning as a case in point (see Marescaux, this volume, for a complementary discussion of the dynamic control tasks).

In the artificial grammar learning task introduced by Reber (1967; see also his 1989; for other reviews see Berry & Dienes, 1993; Dienes & Berry, 1997; Shanks & St John, 1994), a set of rules is used to determine the order that letters can appear in letter strings, which for example may be 5-8 letters long. The rules are sufficiently complex that the ordering of the letters at first seems quite arbitrary to a subject. Subjects are asked to observe, copy, or memorize the letter strings, but they are not told about the existence of the set of rules. After some minutes, the strings are taken away and the subjects are told of the existence of the set of rules, but not what they are. Subjects are asked to classify a new set of strings, half of which obey the rules and half of which do not. The basic finding is that people can classify at above chance rates (typically about 65%) without being able to say freely why they made the decisions they did.

Reber (1967) argued that people had induced rules specifying the structure of the letter strings. His claims about the implicit learning of rules went ignored for a decade or two, but then a flurry of interest started in the 1980's. Dulany, Carlson, and Dewey (1984), Perruchet and Pacteau (1990), and Dienes, Broadbent, and Berry (1991) argued that people had learned allowable small fragments of strings, for example, which bigrams (pairs of letters) occurred in the training strings, and to a lesser extent which trigrams, or higher order n-grams occurred. Such n-gram knowledge could either be learnt as rules which subjects consult explicitly; or as rules that govern subjects' performance but are represented only implicitly. In rule consulting, a rule like "T can follow M" is represented as a fact of the studied strings. Such declarative knowledge would typically be available for reflection on as knowledge. Alternatively, the knowledge may be represented in a fact implicit way, for example in a connectionist network (Dienes, 1992; Dienes, Altmann, & Gao, 1999; Dienes & Perner, 1996). Activation of an M node may lead to activation of the T node via a positive weight; the function of the weight is

[10] Contrast Whittlesea and Dorken's (1993) view that implicit learning is when we learn information for one purpose and do not realize it is relevant for another purpose; the usefulness for the latter purpose is left implicit in the knowledge as that knowledge was originally conceived by the learner. This is a meaning of implicit, but it is different to our meaning, and does not capture the nature of implicit learning as it seems to us: The knowledge can be best suited to the very purpose is was originally acquired for and still be implicit knowledge.

to code the fact that "T can follow M", but this is not explicitly represented as a fact or as knowledge by the weight.

When a subject comes to classify a string the rules the subject has implicitly or explicitly induced about the grammar are used to infer whether the string is grammatical. The subject forms a new piece of knowledge; e.g. "The test string TVXMMM is grammatical". We will call this knowledge a grammaticality judgement, in contrast to the knowledge used to make the judgement (the subject's personal grammatical rules). In order to determine experimentally whether any knowledge is implicit or explicit, we can assess the subject's ability to metacognitively reflect on the knowledge, for example, by asking the person to give a confidence rating on the grammaticality judgement. The situation is analogous to any situation in which a subject makes a metacognitive judgement about their knowledge. For example, consider a subject trying to retrieve the name of a famous person but they have not retrieved an answer yet. How might the person know that they know the answer? Koriat (in press; this volume) distinguished two ways of making such metacognitive judgements: "information-based" in which the person is aware of the inferences they make in forming a conclusion about their knowledge state; and "feeling-based", in which the true inferential basis of the judgements is not explicit, the person is only aware of the result of the inference as a directly-experienced feeling (e.g. the tip of the tongue state). This distinction is the same Rosenthal (2000a,b,c) makes between higher order thoughts that are based on conscious inferences and those that are not. It is only the latter that leads one to be conscious of a mental state and therefore cause a mental state to be a conscious mental state. Thus, if one judges that one knows the unretrieved famous name in a way that appears unmediated (i.e. the tip of the tongue state), one is conscious of the knowledge (but not in respect of all its content) by inferences that are themselves implicit and unconscious. On the other hand, if one judges that one knows the unretrieved famous name because of inferences of which one is conscious ("I suppose I must know the name, because I watched the news quite a bit at that time"), the knowledge of the name does not constitute a conscious mental state, but an implicit one, known about because of explicit, conscious inferences. This interplay between the implicit/explicit nature of the judgement and the inferences leading to the judgement may be part of the reason why different authors have different intuitions on the relation between metacognition and implicit cognition, as discussed by Koriat (see also the chapters in Reder, 1996).

When a subject judges a string to be grammatical, this may be based on inferences that the subject is conscious of as inferences leading to the conclusion that the string is grammatical. The inferential basis of the decision is then explicit. In contrast, implicit learning is a process by which rules are induced about the domain but they are not rules one consults, they are fact implicit, and the person is not conscious of the rules as rules about the domain. When they are applied, the subject is not directly aware of applying knowledge. How could we show experimentally that this was indeed the state of affairs in order to establish the existence of implicit learning?

One way of testing whether people are aware of their grammatical rules is to ask them to describe freely what rules they used. People are bad at describing the knowledge they have induced in an artificial grammar learning experiment (Reber, 1989; Mathews, Buss, Stanley, Blanchard-Fields, Cho, & Druhan, 1989; Dienes, Broadbent, & Berry, 1991). The sceptics however argue that is just because free report is an insensitive test, not because any piece of knowledge is in principle unavailable to free report (e.g. Shanks & St John, 1994). If we knew exactly what type of knowledge structures and rules a subject had, we could ask people to judge whether each rule is one they possess and give a confidence rating to assess the subject's' assessment of their state of knowledge; unfortunately, we can never be quite sure exactly what rules a subject has induced (Marescaux & Chambres, 1999).

Fortunately, we can in principle determine the content and knowledge status of subjects' grammaticality judgements. As implicit learning researchers, the interesting issue is not really the implicit nature of the grammaticality judgements; it is the implicit nature of the

grammatical rules. Nonetheless, the grammaticality judgements can, in certain circumstances, provide a window onto the implicit nature of the grammatical rules. If the grammatical rules are implicit and applied implicitly (i.e. the person does not explicitly represent that they are applying certain rules), the person will be unaware of the inferential basis of their grammaticality judgement. How might the subject decide what confidence rating to give to the grammaticality judgement? The subject may report that the judgement was a pure guess. Thus, if we took all the cases where the subject said the judgement was a pure guess, the implicit knowledge of the grammatical rules would lead the subject to still make correct responses at an above chance (or above control baseline) rate. This is the guessing criterion (Dienes, Altmann, Kwan, & Goode, 1995). The implicit grammaticality knowledge leads to implicit grammaticality judgements (implicit in the sense that the judgement is knowledge to the subject but not represented as knowledge by the subject). In addition, one can examine the relationship between confidence and performance over the whole range of confidence ratings given; this criterion, about to be explained, is called the "zero correlation criterion" (Dienes, Altmann, Kwan, & Goode, 1995).

Reber (reviewed in his 1989) and Dienes, Kurz, Bernhaupt, and Perner (1997) showed that (a) there are different strings to which subjects are differentially consistent in their responding; and (b) subjects are more consistent for strings to which they tend to make the correct response rather than the incorrect response. (a) is evidence that the learning/knowledge application system is treating itself as having different degrees of knowledge about different strings (in responding more consistently to strings it is treating itself as if it had more certain knowledge about those strings than strings it responds inconsistently to); (b) is evidence that it got this correct (as judged by the rules the experimenter had in mind). Are the different knowledge states the subjects treat themselves as being in explicitly represented as such? One can answer this question by determining whether increasing confidence is associated with an increasing tendency to give a correct response. A lack of relationship between confidence and performance (the zero correlation criterion) indicates that subjects do not know that they know; they do not have access to the different knowledge states they are in fact in as being different knowledge states[11]. As reviewed by Dienes and Berry (1997) and Dienes and Perner (in press), this has been found in a number of artificial grammar learning experiments for some types of stimuli (see also Marescaux, this volume, for application to the dynamic control tasks). The implicit and thus unconscious nature of the occurrent knowledge states underlying grammaticality judgements can be taken as a reflection of the implicit nature of the underlying grammar knowledge.

Unfortunately, the window provided by the grammaticality judgements on the implicit nature of the grammatical rules is not always a clear one (e.g. Whittlesea & Dorken, 1997). With practice on the task, subjects may come to base their confidence ratings on cues at least correlated with the knowledge status of the grammaticality judgements (these cues are not known, but could be: reaction times, string fluency, correct explicit knowledge). Indeed, Allwood, Granhag, and Johansson (in press) found that with a task involving a relatively small number of trials, the guessing criterion was satisfied and there was poor calibration of confidence and performance; on a task involving more trials calibration improved dramatically and when subjects claimed they were guessing they were indeed performing at chance[12]. Nonetheless, the result on the latter task leaves open the possibility that the grammar knowledge was quite implicit, given subjects' generally poor ability to freely report the bases of their decisions under conditions very similar to those of Allwood et al. How could one determine the implicit nature of the grammar rules in this situation?

---

[11] Conversely, if (a) and (b) did not jointly hold, the zero correlation criterion would not indicate the existence of implicit knowledge (Dienes & Perner, 1996). If the subject consistently applies a partially correct rule, there is no reason why confidence in correct decisions should be different from confidence in incorrect decisions, regardless of whether the rule is implicit or explicit.

[12] The tasks differed in other respects as well; for example, a different grammar was used. It remains to be determined which factors were responsible for the difference in calibration.

In addition to looking at metacognitive monitoring, one can look at metacognitive control to determine the implicit status of grammar rules. Consistent with the logic of Jacoby (1991), if a subject were asked to complete a string with a letter in such a way that the rules were violated, implicit knowledge of the rules would be hard to inhibit. If the knowledge is applied by rule consulting, it is easy to not apply the rules; one just does not consult them. However, if the knowledge is not represented as knowledge, so it is not represented as the knowledge one has just learnt, there is no means to reference the knowledge by content control in order to inhibit its use. That is, if the knowledge is implicit but activated, the subject will have a tendency to use the knowledge to complete letter strings even when trying not to. This methodology has not been applied to Reber's artificial grammar learning task yet, but Goschke (1998) and Destrebecqz and Cleeremans (in press) applied the methodology to another implicit learning paradigm, the Sequential Reaction Time paradigm, in which a set of rules determine the order in which subjects should press a set of buttons. They found that after training, subjects still pressed the buttons according to the rules despite being told to press in a way that violated the rules. That is, the grammar knowledge could not be brought under content control, and thus its status as a particular body of knowledge must have been implicit. With Reber's artificial grammar learning paradigm, Dienes (1996; Dienes, Altmann, Tunney, & Goode, in preparation) showed that the grammatical status of to-be-ignored flanking strings affected the reaction time of subjects classifying target strings; subjects were faster when the flanking and target strings were consistent rather than inconsistent. That is, the flanking strings were automatically processed for grammatical status; the mere presence of the strings triggered the application of the relevant knowledge schema, indicating the use of vehicle control and thus implicit knowledge.

8. CONCLUSION

This chapter has taken the framework of Dienes and Perner (1999) to show the metacognitive implications of the implicit-explicit distinction in many domains. It has considered the metacognitive basis of the cognitive operations pervasive in all moments of waking life: seeing, perceiving, remembering, willing, and applying knowledge. The simple act of consciously seeing is deeply metacognitive. Understanding the nature of remembering, as much as perception, requires consideration of several layers of metacognition, even putting aside the more obvious metacognitions required to interactively search one's memory store to retrieve an obscure fact of one's past. When a cognitive operation is successful, but metacognition fails, the result is an unconscious mental state. We have argued, based on people's metacognitive failures, that people can perceive unconsciously, strategically act unconsciously, and acquire and apply knowledge of which they are not conscious. An interesting question is raised about the implicit basis of people's poor calibration of metacognitive judgements and their performance in other domains we have not discussed. We hope we can tempt metacognition researchers to look at implicit learning more closely, and implicit learning researchers to look at metacognition more closely.

9. REFERENCES

Allwood, C. M., Granhag, P. A., Johansson, H. (in press). Realism in confidence judgements of performance based on implicit learning. *European Journal of Cognitive Psychology.*

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Armstrong, D. (1980). *The nature of mind and other essays*. Cornell University Press.

Armstrong, D.M. (1991). Intentionality, perception, and causality: reflections on John Searle's *intentionality*. In E. Lepore & R. Van Gulick (Eds.), *John Searle and his critics* (pp. 149-158). Basil Blackwell.

Baddeley, A.D. (1986). *Working memory*. Oxford: Oxford University Press.

Berry, D. C., & Broadbent, D. E. (1984). On the relationhsip between task performance and associated verbalizable knowledge. *Quarterly Journal of Experimental Psychology*, 36A, 209-231.

Berry, D. C. and Dienes, Z. (1993). Implicit learning: Theoretical and empirical issues. Hove: Lawrence Erlbaum.

Brentano, F. von (1874/1970). *Psychology from an empirical standpoint.* (Edited by O.Kraus, translated by L. L. McAllister.) London: Routledge

Bridgeman, B. (1991). Complementary cognitive and motor image processing. In G. Obrecht & L. W. Stark (Eds), *Presbyopia research: From molecular biology to visual adaptation.* Plenum Press.

Broadbent, D. E. (1971). *Decision and stress*. New York: Academic Press.

Broadbent, D. E., & Aston, B. (1978). Human control of a simulated economic system. *Ergonomics*, 21, 1035-1043.

Carruthers, P. (1992). Consciousness and concepts. *Proceedings of the Aristotelian Society, Supplementary Vol. LXVI*, 42-59.

Carruthers, P. (in press). *Phenomenal consciousness naturally*. Cambridge: Cambridge University Press.

Cheesman J. & Merikle, P. M. (1984). Priming with and without awareness. *Perception & Psychophysics*, 36, 387-395.

Cheesman J. & Merikle, P. M. (1986). Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology*, 40, 343-367.

Chrisley, R. L. (1996). *Non-conceptual psychological explanation: Content & computation*. D.Phil. thesis, University of Oxford.

Cussins, A. (1992). Content, embodiment and objectivity: the theory of cognitive trails. *Mind*, 101, 651-688.

Debner, J. A. and Jacoby, L. L.(1994). Unconscious perception: Attention, awareness, and control. *Journal of Experimental Psychology:Learning, Memory, and Cognition*, **20**,304-317.

Destrebecqz, A., & Cleeremans, A. (in press). **C**an sequence learning be implicit? New evidence with the Process Dissociation Procedure. *Psychological Science.*

Dienes, Z. (1987). *Selective attention: Relevance to hypnosis and hypnotizability*. Thesis submitted in partial fulfillment of the requirements for Master of Arts (Hons), School of Behavioural Sciences, Macquarie University.

Dienes, Z. (1992). Connectionist and memory array models of artificial grammar learning. *Cognitive Science, 16*, 41-79.

Dienes, Z. (1996) . How unconscious is the transfer of knowledge of artificial grammars across domains? Invited paper presented at the 1996 Annual Conference of the British Psychological Society, Brighton, 11-14 April, 1996.

Dienes, Z., Altmann, G., & Gao, S-J. (1999) Mapping across domains without feedback: A neural network model of implicit learning. Cognitive Science, 23, 53-82.

Dienes, Z., Altmann, G., Kwan, L, Goode, A. (1995) Unconscious knowledge of artificial grammars is applied strategically. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21*, 1322-1338.

Dienes, Z., & Berry, D. (1997). Implicit learning: below the subjective threshold. Psychonomic Bulletin and Review, 4, 3-23.

Dienes, Z., Broadbent, D. E., & Berry, D. C. (1991). Implicit and explicit knowledge bases in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 17*, 875-882.

Dienes, Z., & Fahey, R. (1995). The role of specific instances in controlling a dynamic system. J*ournal of Experimental Psychology: Learning, Memory, & Cognition*, 21, 848-862.

Dienes, Z., & Fahey, R. (1998) The role of implicit memory in controlling a dynamic system. *Quarterly Journal of Experimental Psychology, 51A*, 593-614.

Dienes, Z., Kurz, A., Bernhaupt, R., & Perner, J. (1997). Application of implicit knowledge: deterministic or probabilistic? *Psychologica Belgica, 37*, 89-112.

Dienes, Z., & Perner, J. (1996) Implicit knowledge in people and connectionist networks. In G. Underwood (Ed), *Implicit cognition* (pp 227-256), Oxford University Press.

Dienes, Z., & Perner, J. (1999) A theory of implicit and explicit knowledge. *Behavioural and Brain Sciences, 22*, 735-755.

Dienes, Z., & Perner, J. (in press) A theory of the implicit nature of implicit learning. In Cleeremans, A., & French, R. (Eds), *Implicit learning and knowledge of the real world*. Psychology Press.

Dokic, J. (1997). *Two metarepresentational theories of episodic memory*. Paper presented at the Annual Meeting of the ESPP in Padua, Italy August 1997.

Dretske, F. (1988). *Explaining behaviour: Reasons in a world of causes.* Cambridge (Massachusetts), London: The MIT Press.

Dulany, D. E. (1996). Consciousness in the explicit (deliberative) and implicit (evocative). In J. D. Cohen & J. W. Schooler (Eds), *Scientific approaches to the study of consciousness* (pp 179-212). Erlbaum: Hillsdale, NJ.

Dulany, D. E., Carlson, R., & Dewey, G. (1984). A case of syntactical learning and judgement: How concrete and how abstract? Journal of Experimental Psychology: General, 113, 541-555.

Flavell, J.H., Flavell, E.R., & Green, F.L. (1983). Development of the appearance-reality distinction. Cognitive Psychology, 15, 95-120.

Gardiner, J. (1988). Functional aspects of recollective experience. *Memory and Cognition*, **16**, 309-313.
Graham, C., & Evans, F. J. (1977). Hypnotizability and the deployment of waking attention. *Journal of Abnormal Psychology, 86, 631-638.*
Goschke, T. (1998). Implicit learning of perceptual and motor sequences: Evidence for independent learning systems. In M. A. Stadler & P. A. French (Eds), *Handbook of implicit learning* (pp 401-444). London: Sage Publications.
Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, 30, 513-541.
Kihlstrom, J. F. (1997). Consciousness and me-ness. In J. D. Cohen & J. W. Schooler (Eds), *Scientific approaches to consciousness.* New Jersey: Erlbaum.
Kihlstrom, J. F., & Cantor, N. (1984). Mental representations of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol 17, pp 2-47). New York: Academic Press.
Kirsch, I (1991). The social learning theory of hypnosis. In Lynn, S. J., & Rhue, J. W. (Eds) *Theories of hypnosis: Current models and perspectives* (pp 439-465). The Guilford Press: New York
Koriat, A. (in press). The feeling of knowing: some metatheoretical implications for consciousness and control. C*onsciousness and Cognition.*
Marescaux, P.-J., & Chambres, P. (1999). What is the cat in complex settings? *Behavioral and Brain Sciences*, 22, 773-774.
Mathews, R.C., Buss, R.R., Stanley, W.B., Blanchard-Fields, F., Cho, J R. & Druhan, B. (1989). The role of implicit and explicit processes in learning from examples: A synergistic effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 1083-1100.
Merikle, P. M., & Joordens, S. (1997). Parallels between perception without attention and perception without awareness . *Consciousness and Cognition*, **6**, 219-236.
Miller, G. (2000). *The mating mind: How sexual choice shaped the evolution of human nature*. London: Heinemann.
Millikan, R. G. (1993). *White queen psychology and other essays for Alice*. Cambridge, MA: Bradford Books/MIT-Press.
Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol 26, pp125-173). San Diego, CA: Academic Press.
Norman, D.A., & Shallice, T. (1986). Attention to Action. Willed and automatic control of behavior.. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation*(Vol.4) (pp. 1-18). New York: Plenum.
Perner, J. (1998). The meta-intentional nature of executive functions and theory of mind. In P. Carruthers and J. Boucher (Eds), Language and thought (pp. 270-283). Cambridge: Cambridge University Press.
Perner, J. (in press, a). Episodic Memory: Essential distinctions and developmental implications. in C. Moore & K. Skene (Eds.). *The self in time: Developmental issues*. Erlbaum. *
Perner, J. (in press, b). Dual control and the causal theory of action: The case of non-intentional action and causal self-reference. In N. Eilan & J. Roessler (Eds), *Agency and self-awareness*. Oxford: Oxford University Press.
Perner, J., & Dienes, Z. (1999). Deconstructing RTK: How to Explicate a Theory of Implicit Knowledge. *Behavioural and Brain Sciences, 22*, 790-808.
Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General*, 119, 264-275.
Perruchet, P., Vinter, A., & Gallego, J. (1997). Implicit learning shapes new conscious percepts and representations. *Psychonomic Bulletin & Review*, 4, 43-48.
Reber, A.S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behaviour, 6,* 855-863.
Reber, A.S. (1989). Implicit learning and tactic knowledge. *Journal of Experimental Psychology: General, 118,* 219-235.
Reder, L. M. (Ed.) (1996). *Implicit memory and metacognition*. Hillsdale, NJ: Erlbaum.
Rosenthal, D.M. (1986). Two concepts of consciousness. *Philosophical Studies*, **49**, 329-359.
Rosenthal, D.M. (2000a).Consciousness, Content, and Metacognitive Judgments, *Consciousness and Cognition*, 9.
Rosenthal, D.M. (2000b). Metacognition and Higher-Order Thoughts (extended response to commentaries), *Consciousness and Cognition*, 9.
Rosenthal, D.M. (2000c).."Consciousness and Metacognition", in Daniel Sperber (Ed.), *Metarepresentation: Proceedings of the Tenth Vancouver Cognitive Science Conference.* New York: Oxford University Press.
Searle, J. R. (1983). *Intentionality*. Cambridge: Cambridge University Press.
Searle, J. R. (1990). Consciousness, explanatory inversion, and cognitive science. *Behavioural and Brain Sciences, 13*, 585-642.
Shanks, D. R. & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioural and Brain Sciences*, 17, 367-448.
Sheehan, P.W., & McConkey, K.M. (1982). *Hypnosis and experience: The exploration of phenomena and process*. Hillsdale, NJ: Erlbaum.

Spanos, N. P. (1986). Hypnotic behaviour: A social-psychological interpretation of amnesia, analgesia, and "trance logic". *The Behavioural and Brain Sciences*, 9, 449-502.

Spanos, N. P., Radtke, H. L., & Dubreuil, D. L. (1982). Episodic and semantic memory in post-hypnotic amnesia: A re-evaluation. *Journal of Personality and Social Psychology*, 43, 565-573.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, **26**, 1 12.

Tye, M. (1995) *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind.* MIT Press: Cambridge.

Tzelgov, J., Ganor, D., & Yehene, V. (1999). Automatic processing results in conscious representations. *Behavioral and Brain Sciences*, 22, 786-787.

Weiskrantz, L. (1988). Some contributions of neuropsychology of vision and memory to the problem of consciousness. In A. J. Marcel & E. Bisiach (Eds.), *Consciousness in contemporary science* (pp. 183-199). Oxford: Clarendon Press.

Whittlesea, B. W. A., & Dorken, M. D. (1993). Incidentally, things in general are particularly determined: An episodic processing account of implicit learning. *Journal of Experimental Psychology: General*, 122, 227-248.

Whittlesea, B. W. A., & Dorken, M. D. (1997). Implicit learning: Indirect, not uncosncious. *Psychonomic Bulletin & Review*, 4, 63-67.