Chapter 19

Seeds of self-knowledge: noetic feelings and metacognition

Jérôme Dokic

Feeling is to knowledge what a cry is to a word.
Erwin Straus

Introduction

As authors from various traditions and disciplines—including phenomenology, cognitive and social psychology—have observed, our most spontaneous judgements can reflect what we ordinarily call 'our feelings'. Sometimes we judge that something is the case just because (*ceteris paribus*) we *feel* that this is so. Feeling-based judgements seem to provide us with information that it would have been difficult, perhaps impossible, to acquire through other epistemic means, such as perception, memory, and inference. As a consequence, they can act as first premises in both theoretical and practical reasoning. In many everyday circumstances, we are ready to judge, reason, and act on the basis of our feelings without further ado.

If ordinary language descriptions of our feelings are to be trusted, the latter can be about external states of affairs ('I feel that it's going to rain'), as well as about our own bodily states and dispositions ('I feel tired', 'I feel elated'). In this chapter, though, I am interested in another species of feelings, namely those that concern our own mental and epistemic life. I shall call the relevant feelings 'noetic feelings'; they have also been called 'epistemic' or 'metacognitive' feelings.[1] Here is a partial and non-exhaustive list of noetic feelings as they have been discussed in the literature:

- *Feelings of knowing/not knowing* (Koriat 1995, 2000).
- *Tip-of-the-tongue experiences* (Brown 2000; Schwarz 2002).
- *Feelings of certainty/uncertainty* (Smith et al. 2003).
- *Feelings of confidence* (Winman and Juslin 2005).
- *Feelings of ease of learning* (Koriat 1997).
- *Feelings of competence* (Bjork and Bjork 1992).
- *Feelings of familiarity* (Whittlesea et al. 2001a, 2001b).
- *Feelings of 'déjà vu'* (Brown 2003).

[1] See Koriat (2006, p. 54), who writes that there is an 'assumption underlying much of the work in metacognition […], that metacognitive feelings play a causal role in affecting judgments and behavior'.

- *Feelings of rationality/irrationality* (James 1879).
- *Feelings of rightness* (Thomson 2008).

These feelings are noetic in the sense that they intuitively concern epistemic states, events, or skills, although the sense in which this is so needs careful delineation. Admittedly, the boundary between noetic feelings and other kinds of feelings is not very sharp. Some feelings seem to lie at the borderline between noetic feelings and feelings about the external world. For instance, it is not clear whether the feeling of presence (Matthen 2005) is just the feeling that a state of affairs is actual (rather than merely possible), or the feeling that one is genuinely *related* to the actual world. Similarly, the feeling that something in the visual field has changed (Rensink 2004; Loussouarn 2010) might really be the feeling that one has *detected* a change, even though one is not able to identify it. In advance of a substantial theory of feelings, it is hard to classify these feelings as genuinely noetic or not. In any case, I shall focus here on feelings which are clearly noetic, such as the feeling of knowing and the feeling of (subjective) uncertainty.

This chapter is structured as follows. In the first section, I discuss a concrete example illustrating the fact that noetic feelings are 'seeds' of self-knowledge, i.e. can provide knowledge or justified beliefs about one's own mental and epistemic life. Then, in the next three sections, I formulate three theoretical models of the psychological nature and epistemic value of noetic feelings. On the Simple Model, noetic feelings are manifestations of metarepresentational states of knowledge that are already in place. On the Direct Access Model, they are (possibly partly opaque) experiences about one's own first-order states of knowledge. Finally, on the Water Diviner Model, they are first and foremost bodily experiences, whose objects (bodily states) are only contingently associated with first-order epistemic states. Still, they can acquire a derived content representing or concerning such states. The latter model will turn out to be superior to the other ones. First, it helps to disambiguate the sense in which noetic feelings can be described as 'metacognitive' ('Metacognition versus metarepresentation' section). Second, it can easily be extended to deal with the motivational dimension that many noetic feelings seem to have ('Noetic feelings and motivation' section). In the following section ('Two kinds of metacognition, and a case study'), I build on the account sketched in the previous sections and illustrate the distinction between two kinds of metacognition (which I call 'procedural' and 'deliberate') with respect to feelings of uncertainty experienced in the context of certain perceptual categorization tasks. Eventually, in the section entitled 'The Competence View', I put forward a tentative hypothesis about the derived intentional contents of noetic feelings, according to which they can concern our own mental and epistemic life without being strictly speaking metarepresentational, i.e. without being constitutively linked to the possession of metarepresentational or mindreading abilities.

Feelings of knowing and self-knowledge

Consider the following pair of questions:

Q1 Is Lima the capital of Peru?

Q2 Do you believe that Lima is the capital of Peru?

On the face of it, these are very different yes–no or polar questions, despite the fact that they have overlapping contents. Q1 is a question about the geographical world, whereas Q2 is a question about the addressee, more precisely about whether she is in a specific mental state, namely the state of believing that Lima is the capital of Peru. Yet the answer to Q2 can be directly based on an answer to Q1. The addressee can answer 'yes' to Q2 if she is ready to answer 'yes' to Q1. Indeed, if she fully understands both questions, she normally *cannot* answer 'yes' to Q2 without thereby being in a position to answer 'yes' to Q1.

Gareth Evans has drawn the connection between these two types of questions in the following general terms:

> I get myself in position to answer the question whether I believe that *p* by putting into operation whatever procedure I have for answering the question whether *p*. (Evans 1982, p. 225.)

In a later essay, Gordon (1995) calls the procedure that Evans is describing here an 'ascent routine':

> Because this procedure answers a metacognitive question by answering a question at the next lower semantic level, I will call it an *ascent routine*. (Gordon 1995, p. 60.)

Both Evans and Gordon take ascent routines to be *alternatives* to traditional introspective methods. In answering Q2, the addressee does not have to search her mind for a specific belief, much less a state of knowledge. Rather, she directs her attention to the outer world as she conceives it. No introspective ability needs to be invoked in order to determine whether she believes that Lima is the capital of Peru.

Now consider another pair of questions:

Q3 What is the capital of Peru?

Q4 Do you know what the capital of Peru is?

Q3 and Q4 are very different non-polar questions, despite the fact that they have overlapping contents. The former is about the geographical world, whereas Q4 is a question about the addressee. Yet the addressee *can* answer Q4 (by saying 'yes') without being in a position to answer Q3 (by saying 'yes, Lima'). In fact, she can answer Q4 without having any city in mind.

There are two ways she can do this. One way is to use independent information to the effect that she is competent in answering a first-order question such as Q3. For instance, she knows that she was a good geography student at school, and that she learnt all the capitals in the world by heart. In such a case, her metacognitive judgement to the effect that she can answer Q3, on which she can ground a 'yes' answer to Q4, is *theory-based*. It inferentially derives from independent beliefs based on memory. Alternatively, the addressee may just *feel* that she knows what the capital of Peru is. She feels competent in answering Q3, in advance of actually providing any answer, either privately or publicly. In this case, her metacognitive judgement is *experience-based*. It seems to be directly based on her affective experience (a 'gut feeling') independently of background beliefs.[2]

What is the nature of the feeling of knowing which enables one to answer a question such as Q4 in advance of giving any answer to Q3? In particular, since ascent routines are clearly not available in this case, is such a feeling a form of introspection of one's own epistemic states? In what follows, I shall present three models of feelings of knowing that try to provide answers to these questions.

The Simple Model

On the Simple Model, noetic feelings are in fact metarepresentational beliefs, more precisely beliefs that are explicitly about one's epistemic states (Dienes and Perner 1999). For instance, the feeling that the subject knows the name of the capital of Peru is just the actualization of a piece of higher-order knowledge ~~that she~~ acquired long ago, namely the knowledge that ~~there is a name such that~~ she knows ~~that it refers to the capital of Lima~~. Of course, if she is wrong and in fact she does not know that the capital of Peru is called 'Lima', her feeling expresses mere apparent

[2] The distinction between theory-based (or information-based) and experience-based metacognitive judgements comes from Koriat (2006).

knowledge, but it is still the actualization of a higher-order mental state, more precisely a false belief about her first-order state of knowledge.

The Simple Model can thus provide a straightforward explanation of why we can have feelings of knowing while being actually unable to retrieve the relevant name, as it happens in so-called 'tip-of-the-tongue' experiences. Surely, the higher-order state of knowledge, or apparent knowledge, that we know the name of the capital of Peru can be made explicit while the corresponding first-order state of knowledge, or apparent knowledge, that the capital of Peru is called 'Lima' remains implicit because of some performance problem. These can be distinct states, and either one can be activated independently of the other. In the case of geographical ignorance, a higher-order state of apparent knowledge that we know the name of the capital of Peru can even exist in the absence of any first-order state of knowledge to the effect that Lima is the capital of Peru.

I call this model 'simple' because it does not posit new kinds of mental states, since noetic feelings are assimilated to ordinary beliefs, in the form of higher-order memory states. On this model, noetic feelings can justify other beliefs because they are themselves beliefs. Besides, we often lose the original justification of our memory beliefs, a fact that might be invoked in order to explain why we are not fully aware of the underlying reasons for what our feelings tell us. Despite its relative simplicity, though, the Simple Model faces several difficulties.

The first difficulty will become clearer as we proceed. It concerns the fact that on the Simple Model, noetic feelings necessarily have metarepresentational contents. They are explicitly about first-order states of knowledge. It follows that the subject must possess relevant epistemic concepts, such as the concept of knowledge or memory, in order to *have* noetic feelings. In order words, noetic feelings are available only to creatures possessing a theory of mind. However, as we shall see (see especially the last two sections), there are reasons to think that creatures lacking metarepresentational resources can still have noetic feelings, such as feelings of knowing and feelings of uncertainty, and exploit them in theoretical and practical reasoning.

Another difficulty is that even if the subject has metarepresentational abilities, noetic feelings seem to be sources of *original* knowledge or justified beliefs, at least in some cases. After all, perhaps the subject never acquired the higher-order knowledge that she knows the name of the capital of Peru, or she might have forgotten about it a long time ago. Still, she can have the feeling that she has such knowledge just because she is being asked a question such as Q3 ('What is the capital of Peru?'). In this case, it seems that her feeling of knowing enables her, in concert with the fact that she possesses the relevant mental concepts, to acquire a new piece of higher-order knowledge. In contrast, if feelings of knowing are already conceived as higher-order beliefs, it is not clear that they can be justified or warranted.

Finally, the Simple Model forces its proponents to adopt a curious interpretation of well-replicated experimental results. It appears that feelings of knowing can be easily manipulated in certain experimental conditions (see, e.g. Reder 1987; Bjork 1999). For instance, by *priming* some of the question terms, psychologists can raise the feeling of familiarity toward a question such as Q3, and produce a fairly convincing feeling that the subject knows the answer to the question, even if she does not. On the Simple Model, these experimental manipulations must be interpreted as creating *false* higher-order memories in the subject, which is quite implausible, at least on the face of it.

The Direct Access Model

On the Direct Access Model, noetic feelings are cases of *introspection*. They provide us with internal awareness of our own first-order memories as carrying information relevant to answering certain questions. So when the subject feels that she knows the name of the capital of Peru, she has

in fact access to one of her first-order states of knowledge, namely the memory that the capital of Peru is called 'Lima'. In the case in point, the subject is not conscious of her memory as having the content 'The capital of Peru is called "Lima"'. Rather, she is conscious of her memory only as having a content of the form 'The capital of Peru is called ____'. In other words, she has introspective access to her memory as such while having access only to a *proper part* of its content.[3] Of course, if the subject does not really know that the capital of Peru is called 'Lima', her feeling of knowing is somehow non-veridical. Still, in this case, she has the *apparent* introspective experience of having the relevant information stored in her mind.

Unlike the Simple Model, the Direct Access Model can explain why noetic feelings are, at least sometimes, a source of original knowledge or justified beliefs about our mental states and dispositions. The subject's feeling of knowing can reveal a piece of information about herself that she may never have explicitly acquired before, namely that she possesses information relevant to answering a question such as Q3.[4] Noetic feelings belong to a class of *experiential* states, so that beliefs based on them can act as bona fide premises in theoretical and practical reasoning. In other words, these beliefs are justified by a belief-independent affective experience, just as perception or memory beliefs are justified by belief-independent perceptual or memory experiences.

It is helpful to compare the Direct Access Model with David Rosenthal's analysis of the tip-of-the-tongue experience:

> When I have Mark Twain's real name on the tip of my tongue, I must be conscious *of* the particular state that carries that information. But I am not conscious of that state in respect of the specific information the state carries; rather, I am conscious of the state only *as* a state that carries that information. (Rosenthal 2000, p. 204.)

Rosenthal draws a distinction between being conscious of a given informational state (for instance, the memory that Mark Twain's real name is 'Samuel Clemens') in respect of the specific information the state carries and being conscious of it only in respect of what questions the information would answer. However, Rosenthal does not defend the Direct Access Model, because he makes clear that being conscious of a given informational state only in respect of what questions the information would answer does *not* entail that this state is itself a conscious state. In contrast, at least to the extent that the objects of conscious introspection must themselves be conscious states, the Direct Access Model entails that feelings of knowing are ways of bringing to consciousness relevant informational states, even though their contents are at least partly occluded to the subject.

Of course the Direct Access Model is hostage to a substantial theory of introspective knowledge, and in particular to the issue of whether the latter should be conceived on the model of observational knowledge. Independently of this issue, though, it is important to notice that the Direct Access Model, at least as applied to feelings of knowing, is incompatible with two general views about introspective knowledge. The first view is that introspection makes the subject aware of her own intentional mental states only by revealing their contents (see, e.g. Tye 2009). In other words,

[3] The Direct Access Model is not committed to the claim that all types of noetic feelings involve opacity in this sense. Certainly feelings of knowing are not unique in this respect. For instance, on this model, the feeling of familiarity relative to a particular perceived person would be the introspective experience of memories involving this person, but whose contents are at least partly opaque to the subject. In other words, the subject knows that she has memories about the person while being temporarily unable to access the full contents of these memories.

[4] So the subject knows that she is competent in answering certain questions *in virtue* of the fact that she is aware of one of her memories as carrying information of a certain kind.

introspection is *fully transparent* with respect to the contents of the introspected states (whenever they have contents). The Direct Access Model denies that introspection is always transparent in this sense, since feelings of knowing are precisely introspective states about particular first-order memories, while their contents are only partially revealed to the subject.

Another, less radical view of introspection or self-knowledge that is incompatible with the Direct Access Model is the 'hierarchy of explicitness' view (as we may call it) according to which the awareness of the contents of one's own mental states is a *precondition* of the awareness of the fact that one is in them (Dienes and Perner 1999, 2002). Unlike the first view, this view acknowledges that introspection can reveal the *mode* of the introspected state, but only if the latter's content has already been fully revealed to consciousness. In contrast, the Direct Access Model allows for a mode to be revealed by introspection (in the case in point, the fact that the introspected state is a *memory*), while only revealing part of the introspected state's content.

At this stage, the Direct Access Model might seem to be a more serious competitor than the Simple Model. Still, the empirical evidence is not quite favourable to it. Psychological experiments suggest that what determines feelings of knowing need *not* be familiarity with the answer. Rather, at least some feelings of knowing are determined by familiarity with question terms (Reder and Ritter 1992) and/or accessibility of partial information regardless of its adequacy (Koriat and Levy-Sadot 2001). In other words, the implicit mechanisms underlying the feeling of knowing need not monitor the memory trace itself (*pace* Hart 1965). In fact, they can be causally disconnected from the subject's first-order state of knowledge. Insofar as the notion of sensitivity is a causal-informational one, they are not sensitive (they do not have direct access) to the presence in long-term memory of the name to be retrieved.[5]

It follows that a natural causal explanation of introspective awareness is not available to proponents of the Direct Access Model. According to this explanation, a necessary condition of being introspectively aware of a mental state M is that M be the *cause* of one's introspective awareness. However, empirical evidence suggests to the contrary that feelings of knowing are not caused by first-order memory states (or corresponding memory traces in the brain), but rather by cues (processing fluency, availability of partial information) that are only contingently associated with these states, which might not even exist. Now whether this is incompatible with the claim that feelings of knowing involve a form of direct introspective access to one's own mental states at the personal level remains to be determined.

The Water Diviner Model

The Water Diviner Model is named after a character introduced by Wittgenstein in *The Blue Book*, who claims to *feel* (the German verb is 'fühlen') in his hand that there is water three feet underground. On this model, noetic feelings are first and foremost bodily experiences, i.e. experiences about bodily states. They are diffuse affective states registering internal physiological conditions and events. Unlike bodily sensations, though, noetic feelings need not have precise locations in external bodily parts. At a phenomenological level, they often have an 'indistinct, spreading, blurred quality' and they 'seem to actively resist attempts to focus attention directly on them' (Mangan 2001). In William James's terms, they belong to the 'fringe of consciousness' (James 1890).

Now, just as the water diviner's sensations reliably co-vary with physical conditions, namely the presence of water underneath, noetic feelings reliably co-vary with *mental* conditions.

[5] Of course, other types of noetic feelings may be such that their underlying metacognitive mechanisms are causally sensitive to the relevant target in memory. Metcalfe (2000) argues that this is the case with 'feelings of imminence', such as those involved in tip-of-the-tongue experiences.

For instance, the feeling of knowing co-varies with the fact that the subject has knowledge about the relevant subject-matter. As a result, at least some particular feelings of knowing indicate or carry information about the presence of first-order states of knowledge. In other words, feelings of knowing 'track' such states, in the sense that normally, the former occur only in the context of the latter ('I would not have the feeling of knowing this person's name if I did not know it'). The cues underlying noetic feelings are contingently but stably associated with epistemic states. This association holds in a normal (ecological) context, but it can be severed by psychologists, who can easily produce 'illusory' feelings of knowing (Bjork 1999).

The informational properties of many token feelings can be exploited by a sophisticated cognitive system to recruit types of feelings as representations of mental states. In other words, there is room for an account of noetic feelings that is analogous to familiar teleological-functionalist accounts of emotions. For instance, Prinz (2004, 2007) argues that emotions are perceptions of bodily states that are recruited to represent core relational themes or concerns, such as danger or loss. In his terminology, the 'nominal' contents of emotions are bodily changes, but the 'real' contents of emotions are core relational themes. Similarly, one may argue that the nominal contents of noetic feelings are bodily changes, but the real contents of feelings are mental states.

However, the analogy between noetic feelings and emotions breaks down at a crucial point. The association between basic emotions and their real contents is robust, and possibly innate. It is difficult to imagine fear that does not have the function of detecting danger. In contrast, many noetic feelings seem to be recruited by the organism through some form of learning.[6] As an illustration, consider Harris et al.'s (1981) findings. Both 8- and 11-year-old children read anomalous sentences in a story more slowly. However, only the older group is able to pick out the anomalous lines as not fitting the story. According to the authors' interpretation, both groups of children generate 'internal signals' of comprehension failures, but only the older children have learned to *exploit* such signals to locate the *source* of their feelings of difficulty.

These results suggest that the *same* type of noetic feelings (in the case in point, feelings of difficulty or easiness), individuated in bodily terms, can have additional, *acquired* contents that can be exploited in judgements.[7] In the case of organisms possessing metarepresentational abilities, these acquired contents can be explicitly about their own mental states. For instance, feelings of knowing can be recruited as feelings *that* one knows something, by deploying the mental concept of knowledge. It remains an open issue whether noetic feelings can have acquired contents that somehow hinge on the presence of mental conditions but *without* representing them as such. (See the following sections for further discussion of this point.)

According to the Water Diviner Model, feelings have intentional contents beyond the body, but only in a derived way, through some kind of learning or association process. Such a process generates new heuristics, i.e. cognitive shortcuts that enable us to move spontaneously from our feelings to judgements concerning the task at hand. One form that such heuristics can take is that of answering for oneself the question 'How do I feel about it?' in order to simplify a task that is

[6] See Proust (2008). I do not deny that non-basic emotions, such as respect or pride, need to be trained. It is an interesting question whether there is anything like the distinction between basic and non-basic emotions in the case of epistemic feelings, but here I shall leave this question open.

[7] Another interpretation of the results is that the younger children lack the feelings that older children have and exploit. But certainly, the former *behave* as if they felt the difficulty of certain passages, which they spontaneously read more slowly. What this suggests is that feelings of difficulty already involve some low-level metacognitive control, which falls short of the ability to exploit these feelings at the level of explicit reasoning.

particularly complex and demanding (Schwartz and Clore 1996).[8] In the specific case of noetic feelings, the relevant heuristics enable the subject to form non-inferential judgements about her own mental states, such as the judgement that she knows how to answer the question she is being asked.[9]

In some cases, the association between noetic feelings and their 'real' contents can be easily broken. According to Reber et al. (2004), the judgement that a picture is likeable can be based, *ceteris paribus*, on positive affect elicited by processing fluency. Now in the experiments of Winkielman and Fazendeiro (in preparation), some participants were informed that factors having nothing to do with the pictures, such as background music, might influence their feelings toward the pictures. These participants actually cease to experience the pictures as likeable (or likeable to the same extent), undermining the connection between positive affect and positive aesthetic judgement.

In other cases, the heuristics underlying the formation of feeling-based judgements are more robust, and might exhibit modularity effects. For instance, I can get the feeling that I know the person in front of me despite of the fact that I independently know (e.g., from reliable testimony) that my feeling is misleading; I do not know this person at all. Still, the cognitive impression that I know her might persist, at least for a while. However, although feelings can be synchronically modular in this sense, depending on the robustness of the relevant heuristics, they are certainly not diachronically modular. It is possible in principle that noetic feelings lose their contents and acquire different ones, as new heuristics are implicitly learned.

Metacognition versus metarepresentation

I have presented three models of the psychological nature and epistemic value of noetic feelings, focusing on the case of feelings of knowing. Even though it is possible that the Simple Model and the Direct Access Model have some validity with respect to particular cases of noetic feelings, the Water Diviner Model seems to have the widest domain of application. It does not face important objections like its competitors, and it is empirically plausible. In general, the intentionality of noetic feelings beyond the body is not intrinsic but derived. Feelings are intrinsically about the body, but some of them—the noetic ones—can be exploited by the subject as more or less reliable symptoms of the instantiation of mental states or conditions.

The Water Diviner Model acknowledges a distinction between the cognitive processes underlying and grounding noetic feelings and the further, independent cognitive processes that enable the subject to exploit noetic feelings in explicit judgement and reasoning. What I wish to show now is that this distinction helps us to disambiguate the common claim that noetic feelings are 'metacognitive'.

[8] Note that the use of these heuristics involves the self-ascription of feelings as such. This is not the general case. We often move directly from our feelings to metacognitive judgements without going through a representation of feelings as such. Moreover, the Water Diviner Model is compatible with the claim that the process of associating bodily states with specific mental states is coeval with the development of new perceptual-recognitional abilities with respect to the former. In other words, bodily experience itself may be enhanced by the association process.

[9] The notion of non-inferentiality at stake here concerns the personal level. Feeling-based judgements are cognitively spontaneous in something like Bonjour's sense, i.e. they are involuntary, 'coercive,' and not the result of any *introspectible* train of reasoning (Bonjour 1985, p. 117). Of course this is compatible with their being based on subpersonal inferences or computations.

Psychologists usually define metacognition as 'cognition about one's own cognition', or as 'thinking about thinking'.[10] Philosophers, on the other hand, tend to equate metacognition with metarepresentation, i.e. the ability to form representations about other representations, which is usually associated with possessing a mindreading ability or 'theory of mind'.[11] Correspondingly, contents are metarepresentational when they are explicitly about representations as such. For instance, the content of the belief that Pierre believes that it is going to rain is metarepresentational, because of the presence in it of the mental state of *believing* that it is going to rain.

In fact, noetic feelings can be said to be metacognitive in two quite different senses, depending on whether we are talking about their consciously experienced *intentional contents* or their implicit *causal antecedents*.

Firstly, noetic feelings can be said to be metacognitive insofar as their intentional contents yield information (or misinformation) concerning one's own epistemic states, processes, and abilities. The question is whether these contents are also metarepresentational, which would entail that their apprehension required the possession of mindreading abilities. Here we face two alternatives. If we answer 'yes', no creature can exploit noetic feelings in reasoning without deploying some mental concept or proto-concept. For instance, the content of the feeling of knowing a person's name can only be as sophisticated as *that I know this person's name*, which is the representation of a knowledge state as such. In contrast, if we answer 'no', we allow for the possibility that noetic feelings can rationally guide decision-making and the fixation of beliefs in creatures lacking metarepresentational abilities. Of course, the challenge faced by the second alternative is to show that noetic feelings can be self-directed while having first-order contents. As we shall see in a later section ('Two kinds of metacognition, and a case study'), this challenge is highly relevant to the issue of the correct interpretation of important results in the field of animal cognition.

Secondly, the causal antecedents of noetic feelings can be said to be metacognitive insofar as they involve implicit *monitoring* mechanisms that are sensitive to non-intentional properties of first-order cognitive processes. For instance, the feeling of knowing can be based on an implicit evaluation of the *fluency* of the process constituting our spontaneous attempt to remember something. The feeling of familiarity seems to be based on the implicit detection of a discrepancy between expected and actual fluency of processing (Whittlesea et al. 2001a, 2001b). Obviously, the operations of these mechanisms do not require metarepresentational abilities. To begin with, they are sensitive to properties of internal states and processes independently of whatever *contents* they are carrying. If they involve representations of other representations, they do not involve metarepresentations, i.e. representations of representations *as of* representations.[12]

There may be another, more controversial consideration that leads to scepticism about the possibility that implicit metacognitive mechanisms manipulate metarepresentations. One might argue that metarepresentations are necessarily either actually or potentially conscious. There is a constitutive link between the ability to form metarepresentations and the ability to enjoy conscious states. Metarepresentations involve some conception of mental representation, whose complexity makes them available only to conscious creatures and not to sub-personal mechanisms. In contrast, implicit metacognitive mechanisms involve only representations, which

[10] See, for instance, Nelson (1992) and Metcalfe and Shimamura (1994).

[11] A notable exception is Proust (2006, 2007, 2008), who has forcefully and convincingly argued that metacognitive abilities are distinct and independent from metarepresentational abilities.

[12] As Koriat puts it, judgements based on feelings of knowing 'rely on *contentless* mnemonic cues that pertain to the quality of processing, in particular, the fluency with which information is encoded and retrieved' (Koriat 2006, pp. 19–20; my italics).

cannot be or become conscious. As a consequence, they cannot be metarepresentations. They are first-order representations happening to be about internal rather than external states. In a nutshell, they are first-order but self-directed, as opposed to world-directed.

The two senses in which noetic feelings involve metacognitive abilities are largely independent from each other. Even if one acknowledges that the causal antecedents of noetic feelings involve mechanisms that are implicitly sensitive to the quality of first-order processes, the question of whether the intentional contents of noetic feelings can be metacognitive without being metarepresentational remains entirely open. (We shall come back to this question in the section entitled 'The Competence View'.)

Noetic feelings and motivation

Even if the Water Diviner Model is on the right track, it is still incomplete in that it does not deal with an important feature of many types of noetic feelings, namely their *motivational* dimension. Unlike mere intuitions, noetic feelings can intrinsically motivate the subject *to do* something, either at the mental level (e.g., to form a *judgement*) or at the physical level (e.g., to issue a *speech-act* in order to answer a question).[13]

Consider, for instance, tip-of-the-tongue experiences. They are at least partly constituted by a spontaneous *inclination* or *tendency* to search one's memory and retrieve the relevant information (e.g. the proper name that one has on the tip of ~~our~~ tongue). It is hard to imagine having a tip-of-the-tongue experience in the absence of such inclination. Of course, one may be independently motivated, at a higher level, not to waste too much time on the task at hand, but it may be hard to resist the primitive inclinations provided at a lower level by one's feeling of knowing. Noetic feelings have a quasi-modular motivational dimension, analogous to the quasi-modularity of emotions (de Sousa 1987).

One may hypothesize that the motivational power of noetic feelings *derives* from their causal antecedents, which involve mental events of *trying* to do something. In other words, noetic feelings piggyback on intrinsically motivational states that already fix a (mental and/or physical) goal for the subject.[14]

This hypothesis highlights the Janus-faced character of noetic feelings with respect to behaviour. Noetic feelings both *precede* and *follow* behaviour. On the one hand, noetic feelings precede and causally determine actions, by providing first premises to practical reasoning. For instance, we can exploit a feeling of incompetence relative to a particular test in a practical deliberation over whether we should take the test or not. Let us call 'Type 2' the controlled, deliberate behaviour that can be initiated by noetic feelings. On the other hand, noetic feelings follow or at least accompany inclinations to act that are already in place. For instance, psychological experiments have revealed that the feeling of knowing a person's name can be based on the unconscious feedback from the subject's spontaneous attempt to retrieve the name from memory. We feel that we know the name of the person we are talking to because we are already *trying* to remember it, and perhaps retrieving at least part of the relevant information (such as the fact that the name is

[13] I do not want to claim that all types of noetic feelings have a motivational dimension. For instance, perhaps 'déjà vu' experiences are independent of any inclination to act, physically or mentally.

[14] I assume that the relation between noetic feelings and antecedent behaviour is *causal*, and thus contingent. A stronger assumption is that this relation can be at least partly *constitutive*. On this assumption, at least some noetic feelings *are* in fact bodily facets of tryings.

dissyllabic), even though we cannot consciously access the whole of it.[15] We can call 'Type 1' the spontaneous behaviour that gives rise to noetic feelings.[16]

The fact that noetic feelings follow behaviour is congenial to an analysis of feelings along the lines of the James–Lange theory of emotions (Koriat et al. 2006; Laird 2007). According to this theory, which James contrasted with the commonsensical view that emotions cause behaviour, 'we feel sorry *because* we cry, angry *because* we strike, afraid *because* we tremble' (James 1890, p. 449). When transposed to noetic feelings, the claim is that we have a feeling of knowing *because* we are already trying to retrieve the relevant piece of information (Type 1 behaviour). However, unlike what James assumed in the case of emotions, this claim need not be in conflict with common sense insofar as feelings can also be the starting point of further, Type 2 behaviour.

The motivational character of tryings underlying noetic feelings *constrains* the intentional content of the latter as it is exploited in conscious reasoning. For instance, the feeling of knowing (respectively, the feeling of *not* knowing) is causally based on the subject's trying to remember the name, and partly determines the strategies that should be deployed at the level of practical reasoning, by providing information (or misinformation) to the effect that the relevant name can be found in the subject (respectively, elsewhere, in other more competent persons or in a book). Such pre-established harmony is no mystery as soon as we acknowledge the stepwise character of noetic feelings. It also shows that the derived intentionality of noetic feelings is not as arbitrary as, say, the derived intentionality of language. One cannot interpret noetic feelings in any way we like, on pain of creating behavioural dissonance.

Two kinds of metacognition, and a case study

Let's take stock. What has emerged from the previous two sections is a general distinction between two kinds of metacognition, which I will henceforth call 'procedural' and 'deliberate'. *Procedural metacognition* is constituted by implicit monitoring and control of first-order processes. Procedural metacognition can generate conscious feelings, but the latter remain epiphenomenal in the sense that they do not mediate the interactions between monitoring and control. Feelings are neither causal nor epistemic intermediaries in the processes of procedural metacognition. At the personal level, procedural metacognition appears as a purely practical skill, which manipulates only implicit representations.[17]

Procedural metacognition can be contrasted with *deliberate metacognition*, which enables the rational exploitation of noetic feelings. There is deliberate metacognition when noetic feelings give rise to judgements that can be used in practical and theoretical reasoning. Deliberate metacognition is something that the subject herself does, rather than a mechanism inside her. As we have seen, the question arises whether deliberate metacognition involves metarepresentational

[15] See, for instance, Koriat and Levy-Sadot (2000), Koriat (2006), and Koriat et al. (2006). As Koriat (1995, p. 312) writes: 'It is by attempting to search for the solicited target that one can judge the likelihood that the target resides in memory and is worth continuing to search for'.

[16] The Type 1/Type 2 terminology is of course reminiscent of the System 1/System 2 distinction, which has been used to characterize two systems of reasoning, intuitive and deliberate (see Kahneman and Frederick 2005; Evans and Frankish 2008). However, if Type 2 behaviour is indeed deliberate, I want to leave open here whether Type 1 behaviour necessarily belongs to System 1—perhaps there is also something like monitoring targeted at processes belonging to System 2.

[17] See Reder and Shunn (1996) and Spehn and Reder (2000) for further discussion of the claim that metacognitive monitoring and control need not be mediated by conscious awareness.

abilities or not. So there is in principle a further distinction between two species of deliberate metacognition, one which involves metarepresentations and the other which does not.

A difficult question is whether noetic feelings are *necessarily* based on procedural metacognition. Clearly, many noetic feelings result from the feedback from implicit control processes (Koriat et al. 2006), which are instances of procedural metacognition in the sense just introduced. One might still wonder whether some noetic feelings result from a *dedicated* form of monitoring, i.e. one that enables control only at the conscious, rational level. Although this is not a priori inconsistent, it is empirically doubtful. Given the brain's ability to create cognitive shortcuts, one can surmise that once such a monitoring mechanism is in place, its outputs will soon be exploited directly at the subpersonal level, without the mediation of conscious experience. Thus, it seems to be an empirical fact that deliberate metacognition (whether it takes a metarepresentational form or not) is always based on procedural metacognition, and thus that noetic feelings are essentially motivational in the sense that they reflect behavioural inclinations that are already in place.

In the rest of this section, I would like to apply the distinction between procedural and deliberate metacognition to a case study that comparative psychologists have recently set up. This case study is about another type of noetic feelings, namely feelings of uncertainty as they can arise in some perceptual categorization tasks. Hopefully this will also illustrate the relevance of the distinction for a general theory of noetic feelings.

It has been argued that at least some non-human animals, including dolphins and some species of monkeys, have noetic feelings, such as feelings of uncertainty, which they can use strategically in their reasoning (Smith et al. 2003; Smith 2005, 2009). For instance, in one of David Smith's numerous experiments, a monkey has to touch a visual pattern on the screen when it is judged to be dense, and the symbol 'S' when the pattern is judged to be sparse instead. In another condition, the monkey is also allowed to press a third, so-called 'uncertainty' key, which simply advances it to the next trial. Like human subjects, the monkey can make an adaptive use of the uncertainty key by reducing the number of errors that it would make in a forced-choice condition. Moreover, it uses this key in conditions very similar to those in which human subjects verbally report that they *felt unsure* about the category of the stimulus. Now if monkeys can have feelings of uncertainty, they should have first-order contents, since most present-day researchers are reluctant to grant non-human animals full-fledged metarepresentational abilities.[18]

Carruthers (2008, see also 2009) speculates about the mechanism underlying feelings of uncertainty in such cases, which he calls 'the gate-keeping mechanism': 'when confronted with conflicting plans that are too close to one another in strength [it] will refrain from acting on the one that happens to be strongest at that moment, and will initiate alternative information-gathering behaviour instead' (Carruthers 2008, p. 66). The gate-keeping mechanism operates when different goals are competing with one another to control behaviour. It initiates one of the desired behaviours only if the desires involved are not too close to one another in strength. For instance, because of the ambiguity of his visual categorizations, the subject is both weakly inclined to press the 'dense' key, and weakly inclined to press the 'sparse' key. Carruthers points out that the gate-keeping mechanism deals with the fact that 'perceptual processes are inherently noisy' (Carruthers 2008, p. 67). No two perceptual beliefs will have the same strength even given the same stimuli. Correspondingly, the subject's inclinations to act won't be stable over time, even ~~though~~ the world itself does not change.

Carruthers makes clear that the operations of the gate-keeping mechanism do not require metarepresentational abilities. This mechanism 'is sensitive to one *property* of desire (strength) without needing to represent that it is a *desire* that has that property' (Carruthers 2008, p. 67).

[18] See, for instance, Tomasello (1999) and Tomasello et al. (2005).

It is causally sensitive to non-intentional properties of first-order mental states, namely the strength that the subject's desires have independently of their contents.

Carruthers gives a more detailed account of the way feelings of uncertainty arise out of the operations of the gate-keeping mechanism. He suggests that they consist in 'an awareness of a distinctive profile of physiological behavioural reactions caused by the activation of the gate-keeping mechanism (including hesitating and engaging in a variety of information-seeking behaviours, such as squinting at the display or looking closer), which is experienced as aversive' (2008, p. 68). In other words, feelings of uncertainty are bodily feelings akin to aversive anxiety. They have first-order contents, insofar as they are about a kind of non-mental, bodily state.

As it stands, Carruthers' account is congenial to the Water Diviner Model and what we have said about the causal origins of noetic feelings. Feelings of uncertainty are bodily feelings that co-vary with states of uncertainty (bodily hesitations, facial tensions, etc.), as they are detected by the gate-keeping mechanism. However, his account neglects the complexity of the relationship among the gate-keeping mechanism, feelings of uncertainty, and behaviour. He seems to treat on a par all behaviours caused by states of uncertainty, whether they are of Type 1 or Type 2. His list of relevant behaviours includes 'hesitating', 'squinting at the display', 'looking closer' (Type 1), but also 'engaging in information-seeking behaviour', 'searching for another alternative' (Type 2). Obviously, 'searching for another alternative' is a highly abstract goal, which cannot be achieved by simple, pre-wired connections between states of uncertainty and behaviour. Rather, what counts as information-gathering behaviour depends on the subject's background beliefs, and hence is a highly contextualized matter.

As we have seen, the role of epistemic feelings in both types of behaviour is very different. On the one hand, implicit metacognitive processes can give rise to spontaneous simple behaviours such as pausing, squinting, moving one's head from side to side, etc. In such cases, which involve forms of procedural metacognition, conscious feelings of uncertainty are epiphenomenal; they do not *intervene* between states of uncertainty and behaviour. On the other hand, these feelings can give rise to new premises participating in further, explicit reasoning. In the latter cases, which involve forms of deliberate metacognition, feelings of uncertainty essentially intervene between states of uncertainty and more controlled behaviour.

So the situation with respect to Smith's non-human animals is more complex than Carruthers seems to suppose. There are in fact three main interpretations of Smith's results:

1. The animals have acquired a new form of procedural metacognition (a new practical skill), but they lack deliberate metacognition. If they have feelings of uncertainty, the latter are epiphenomenal and are not used in explicit practical reasoning.
2. The animals have acquired new forms of both procedural and deliberate metacognition. They can use feelings of uncertainty in explicit practical reasoning without bringing to bear metarepresentational resources (which they lack).
3. The animals have acquired new forms of both procedural and deliberate metacognition. They can use feelings of uncertainty in explicit practical reasoning as having metarepresentational contents (what they feel is that they are *unsure* about their perceptual categorizations).

What would constitute empirical evidence in favour of the animals manifesting deliberate, and not merely procedural, metacognition? Like the other types of noetic feelings, feelings of uncertainty can play an epistemic role in practical reasoning only if they can be 'at the service of many distinct projects', and their 'influence on any project [is] mediated by other beliefs', to borrow the terms used by Gareth Evans in order to characterize the distinction between explicit beliefs and implicit representations (Evans 1985, p. 337). In general, the ability to use noetic feelings as first premises in theoretical and practical reasoning requires a certain degree of *cognitive flexibility*.

Thus, the empirical hypothesis that some non-human animals can make an adaptive use of the 'uncertainty' response turns on the question of whether their behaviour has enough cognitive flexibility. In other words, the question is whether the animals' behaviour when they choose the 'uncertainty' response is spontaneous or deliberate, i.e. rationally mediated by other beliefs. This question cannot be answered just by observing a single piece of behaviour, or the same type of behaviour within a single task. Much more relevant is the finding that an animal has the ability to *transfer* (without new learning) the choice of the 'uncertainty' response across quite different tasks.[19] For this ability indicates a fair amount of cognitive flexibility, which confirms the deliberate character of the animal's response.

If, on the contrary, the animal learns to use the opt-out button but is unable to transfer its competence to other tasks, then we should say that what it acquired is merely a new procedural skill, an original piece of know-how. It knows how to use the opt-out button in a limited class of contexts, in which the same task or very similar ones are at stake. The animal's skill is still metacognitive, but only in the procedural sense. If the animal experiences noetic feelings, the latter are epiphenomenal and play no causal or epistemic role in the animal's behaviour.[20]

Assuming that the animals have acquired a genuine form of deliberate metacognition, how should we arbitrate between the second and the third interpretations? It is an open question whether cognitive flexibility, which arguably can be observed in the animal realm, requires a form of reflexivity, which some consider to be unique to humans. Of course, the kind of reflexivity that is associated with the possession of metarepresentational abilities enables a strong form of cognitive flexibility, but there may be non-reflexive forms of cognitive flexibility as well.

If room is made for the second interpretation, then Smith's results cannot be used to show that non-human animals, such as some species of monkeys, have metarepresentational abilities (and indeed Smith himself does not favour the third interpretation of his results). For these results would be compatible with the fact that noetic feelings have first-order intentional contents. However, what such contents might be has not been determined yet, and to this question I now turn.

The Competence View

In this section, I shall sketch an abstract account of the intentional contents of at least some noetic feelings, which I argue makes them first-order rather than metarepresentational. I shall call this account 'the Competence View'.

A possible strategy would be to suggest that what appears to be metarepresentational information carried by the intentional content of a noetic feeling is in fact carried at the level of its psychological *mode*. For instance, the content of the feeling of uncertainty relative to the state of affairs that *p* is not *that I feel uncertain that p*, but simply *p* itself. The relevant attitude is feeling-uncertain(*p*) rather than feeling(uncertain that *p*). My main worry with this suggestion is that it does not explain what premises feelings of uncertainty add to explicit reasoning. Of course it cannot be the premise that *p* itself. In other words, what needs to be explained is how the contents of judgements spontaneously based on noetic feelings, which correspond to the latter's acquired or 'real' contents, can fall short of being metarepresentational.

[19] See Proust (2006).

[20] Admittedly, if the concept of cognitive flexibility is vague, it will be difficult to draw the boundary between cases in which metacognition is purely procedural and cases in which it involves noetic feelings that yield first premises as a basis for reasoning to a practical conclusion.

According to the Competence View, a particular noetic feeling is about one's own cognitive competence at a given task. Its content can have the form *I can do this* (or the selfless form *This can be done*), where the demonstrative 'this' refers to a relevant cognitive task in the subject's current situation. In this respect, noetic feelings are akin to feelings of physical competence. When I walk down a rocky hill, my readiness to jump from one rock to another may be based on the feeling *that I can do it*. My feeling is about my competence in a *physical* task, namely jumping to a particular rock. What differentiates cognitive from physical tasks is a difficult question. As a first approximation, one can say that success in doing a cognitive task hangs on possessing beliefs or pieces of information that are not immediately transparent in the subject's situation. For instance, solving the bat-and-ball puzzle is a cognitive task because it requires that one *work out* the correct answer (even at the implicit level), which is not immediately given in the puzzle itself.[21]

On the Competence View, noetic feelings provide their subjects with a type of *modal* knowledge. They yield information about what might *easily* happen, now or in the near future. Something might easily happen if it is the case in nearby possible worlds (where the notion of modal proximity is context-dependent). For instance, the feeling of knowing is the feeling that one's performance is or will be successful in possible worlds close to the actual world. Now these worlds can be more or less close to the actual world, depending on the robustness of one's competence. The more robust one's competence is, the less easily one's performance might fail. If one's competence is fragile, one's performance might fail in possible worlds not too distant from the actual one. One might suggest that *degrees* of noetic feelings can then be modelled in terms of the modal extent to which one's performance is successful. A strong feeling of knowing indicates that one should not expect one's performance to fail too easily. In contrast, a weak feeling of knowing indicates that while one can still do the task, one's performance might more easily fail. In short, thanks to their noetic feelings, subjects have some information about the degree of proximity of the worlds in which their performance would succeed or fail.

The Competence View makes noetic feelings first-order *only if* one can represent one's own cognitive competence without representing it as involving beliefs or other intrinsically contentful states. The challenge is to show that the explicit target of noetic feelings is a particular task rather than the beliefs that are required to deal with it. For instance, the feeling of knowing can be the feeling that one *can* answer the question, rather than the feeling that one *knows* the answer to the question—although it is always possible (and perhaps inevitable) for adult human beings to re-describe their feelings in explicitly metarepresentational terms.[22]

However, it does not follow that all rational uses of feelings of certainty and uncertainty require metarepresentational abilities. In general, according to the Competence View, the contents of noetic feelings can be action-oriented rather than belief-oriented. They can tell the subject something about what she is doing or is inclined to do.[23] For instance, feelings of certainty in the

[21] Here is the puzzle: 'A bat and a ball cost $1.10 in total. The bat costs $1 more than the ball. How much does the ball cost?' Many people answer '10 cents'. Kahneman and Frederick (2005, p. 273) comment that 'the surprisingly high rate of errors in this easy problem illustrates how lightly system 2 [the deliberate system] monitors the output of system 1 [the intuitive system]: people are often content to trust a plausible judgment that quickly comes to mind. (The correct answer, by the way, is 5 cents.)'

[22] On the uniquely human tendency to re-describe in metarepresentational terms what are in fact first-order states and processes, see Povinelli (2003). When an initially first-order state is systematically re-described in metarepresentational terms, it may end up *acquiring* a metarepresentational content. Perhaps this is the case with feelings of knowing experienced by human adults.

[23] Then one might object that they are about one's performance rather than one's competence. Assessing one's competence is based on some concept of competence, whereas assessing one's performance is

context of a categorization task may tell the subject something like: 'If you press the 'dense' key, you are guaranteed to be successful'. In contrast, feelings of uncertainty may tell something like: 'Any success in pressing the 'dense' key will be accidental'. In a nutshell, these feelings can have contents of the form 'I can (cannot) succeed in pressing the right key'. This will be the case when what is at stake is one's success in doing a particular task rather than, more specifically, the truth of one's perceptual beliefs, even if the former actually depends on the latter.

Contents of the form 'I can do it' are not metarepresentational, at least in the sense in which contents of the form 'I believe/know that *p*' are metarepresentational. They are modal contents, which presumably entails that their grasping requires some understanding of counterfactual representations. What their grasping does not require, at least when they are used strategically in the context of practical tasks, is the ability to form representations about mental representations, i.e. to have a theory of mind.

It might be objected that even contents of the form 'I can do it' are in fact concealed metarepresentations. David Lewis notes that 'the 'can' and 'must' of ordinary language do not often express absolute ('logical' or 'metaphysical') modality. Usually they express various relative modalities' (Lewis 1983, p. 246), for instance, modalities relative to our stock of knowledge. This is also the case with the notion of competence that is expressed here by the modal verb 'can'. Noetic feelings can tell the subject something about her performance in nearby possible worlds, but what counts as a nearby world is relative to the subject's cognitive abilities, for instance the acuity of her perceptual discriminations. It does not follow, though, that noetic feelings are necessarily *about* one's cognitive abilities as such. One can be aware of a relative property without representing what the property is relative to. For instance, even if colour properties are relative to the structure of our visual system, our colour experiences do not represent our visual system as such.

Conclusion

This essay was about the psychological nature of noetic feelings. I have argued that noetic feelings are neither higher-order beliefs or memories (contra the Simple Model) nor introspective experiences ~~about~~ first-order epistemic states (contra the Direct Access Model). Rather, they are first-order bodily experiences, namely non-sensory affective experiences about bodily states, which given our brain architecture co-vary with first-order epistemic states, in such a way that they can be recruited, through some kind of learning or association process, to represent conditions hinging on relevant epistemic properties of one's own mind. This is what I have called 'the Water Diviner Model'.

Within this model, noetic feelings can be seen to be associated with two kinds of metacognitive abilities, which I called 'procedural' and 'deliberate'. At the procedural level, our brain realizes mechanisms whose function is to monitor the quality of our cognitive processes in order to produce spontaneous mental and/or physical behaviour (such as attempting to remember a name, reading more slowly, or moving one's head from side to side to resolve visual ambiguity). At the deliberate level, the same mechanisms can generate conscious noetic feelings, which can be further exploited in controlled reasoning to produce more context-sensitive behaviour (such as

merely based on trying to do something. However, this objection neglects the *modal* component that feelings of knowing have according to the Competence View. This is where some concept of competence (embodied in the 'can' of 'I can do it') enters the picture. Thanks to Joëlle Proust for prompting me to clarify this point.

going through the alphabet to provoke remembering, pointing to difficult sentences, or using a magnifying glass).

It follows that the question of the relationship between metacognition and metarepresentation divides into two, depending on whether procedural or deliberate metacognition is at stake. On the one hand, procedural metacognition does not require metarepresentational abilities at all, because it does not manipulate representations as of other representations. On the other hand, there is a genuine issue as to whether the (acquired) intentional contents of noetic feelings can be first-order or must be metarepresentational. One might claim that because noetic feelings track epistemic states, their contents can only be explicitly *about* them. However, the fact that subjects discriminate between knowledge and ignorance shows at best that they know *when* they know (at least sometimes), but not necessarily *that* they know. I have tentatively suggested a way of construing the contents of at least some noetic feelings, as being about one's own cognitive competence at a given task, which does not obviously tie them to metarepresentational abilities.

Of course, much more has to be said about the epistemology of noetic feelings. It is generally agreed that noetic feelings are fallible but reliable. Intuitively, though, they are not on a par with perceptual experiences, which have the property of disclosing part of the world to us. It would be odd to suggest that we can *perceive* (even amodally) our likely success in some cognitive task, in the same way that we can visually experience the presence of coffee in the cup. There may be an interesting difference between feelings of cognitive competence and feelings of *physical* competence. We are less reluctant to acknowledge that we can more or less directly perceive our own physical competence in a particular context. For instance, I can be *visually aware* that I can jump to this rock, even if (*pace* J. J. Gibson and his theory of affordances) my perception of my physical competence in this context may not be as direct as my perception of the rock itself. Nonetheless, noetic feelings merely raise the probability that their contents are true, inviting the subject to take them into account in her reasoning. They are metacognitive signals with a significant yet limited epistemic value, at least in comparison with genuine perceptual experiences. This point is no doubt connected to the fact that the contents of noetic feelings, insofar as they concern the subject's own mental and epistemic life, are acquired or derived, in contrast with the intrinsic contents of perception.

Because my interest in this essay was in the relationship between noetic feelings and metacognitive judgements, I have assumed that noetic feelings are conscious, more precisely that they have an essentially conscious aspect. Indeed, the phenomenological observation that noetic feelings belong to the 'fringe' of consciousness is congenial to Koriat's (2006) 'crossover model', according to which noetic feelings lie at the interface between implicit and explicit processes. In contrast, de Sousa (2008) suggests that feelings differ from full-fledged emotions in that they can be 'attributed at a subpersonal level'. However, perhaps there is no real disagreement here. If de Sousa suggests that metacognitive abilities can operate below the level of consciousness, I agree with him, since I have also acknowledged the existence of a procedural form of metacognition. Now de Sousa's suggestion might be interpreted as the claim that procedural metacognition involves nonconscious noetic feelings. Since I am not sure that this claim has any real explanatory bite, I am tempted to think that my disagreement with de Sousa is purely terminological. What is important is the fact that if procedural metacognition involves *conscious* feelings, the latter are epiphenomenal and do not intervene in the implicit dynamics of monitoring and control processes at the subpersonal level.

Acknowledgements

I would like to thank Joëlle Proust and the participants of her 'Metacognition' seminar at the Jean-Nicod Institute in Paris for many insightful comments on an earlier draft of this chapter.

I borrow the metaphor of noetic feelings as 'seeds' of self-knowledge from Alston's classical essay on feelings (Alston 1969).

References

Alston, W. P. (1969). Feelings. *The Philosophical Review*, 78(1), 3–34.

Bjork, R. A. (1999). Assessing our own competence: Heuristics and illusions. In D. Gopher and A. Koriat (Eds.) *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application*, pp. 435–59. Cambridge, MA: MIT Press.

Bjork, R. A. and Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. F. Healy, S. M. Kosslyn and R. M. Shiffrin (Eds.) *Essays in honor of William K. Estes (Vol. 1: From learning theory to connectionist theory; Vol. 2: From learning processes to cognitive processes)*, pp. 35–67. Hillsdale, NJ: Erlbaum.

Bonjour, L. (1985). *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press.

Brown, A. S. (2000). Tip-of-the-tongue phenomena: An introductory phenomenological analysis. *Consciousness and Cognition*, 9(4), 516–37.

Brown, A. S. (2003). A review of the déjà vu experience. *Psychological Bulletin*, 129, 394–413.

Carruthers, P. (2005). *Consciousness. Essays from a Higher-Order Perspective*. Oxford: Clarendon Press.

Carruthers, P. (2006). *The architecture of the mind*. Oxford: Clarendon Press.

Carruthers, P. (2008). Meta-cognition in Animals: A Skeptical Look. *Mind and Language*, 23(1), 58–89.

Carruthers, P. (2009). How we know our own minds: The relationship between mindreading and metacognition. *Behavioral and Brain Sciences*, 32(2), 1–62.

de Sousa, R. (1987). *The Rationality of Emotion*. Cambridge, MA: MIT Press.

de Sousa, R. (2008). Inference and epistemic feelings. In G. Brun, U. Doguoglu, and D. Kuenzle (Eds.) *Epistemology and emotions*, pp. 185–204. Aldershot: Ashgate.

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

Dienes, Z. and Perner, J. (2002). The metacognitive implications of the implicit-explicit distinction. In P. Chambres, M. Izaute, and P.-J. Marescaux (Eds.) *Metacognition: Process, function, and use*, pp. 171–190. Dordrecht: Kluwer Academic Publishers.

Evans, G. (1982). *The Varieties of Reference*. Oxford: Clarendon Press.

Evans, G. (1985). Semantic theory and tacit knowledge. In G. Evans (Ed.) *Collected Papers*. Oxford: Clarendon Press.

Evans, J. St. B. T. (2003). In two minds: dual-process accounts of reasoning. *Trends in Cognitive Sciences*, 7(10), 454–9.

Evans, J. St. B. T. (2007). On the resolution of conflict in dual process theories of reasoning. *Thinking & Reasoning*, 13(4), 321–39.

Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59, 255–78.

Evans, J. St. B. T. and Frankish, K. (Eds.) *In two minds: Dual processes and beyond*. Oxford: Oxford University Press.

Gordon, R. (1995). Simulation without introspection or inference from me to you. In M. Davies and T. Stone (Eds.) *Mental Simulation*. Oxford: Blackwell.

Harris, P. L., Kruithof, A., Meerum Terwogt M., and Visser, T. (1981). Children's detection and awareness of textual anomaly. *Journal of Experimental Child Psychology*, 31, 212–30.

Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology*, 56, 208–16.

Hookway, C. (2003). Affective states and epistemic immediacy. *Metaphilosophy*, 34, 78–96.

Hookway, C. (2008). Epistemic immediacy, doubt and anxiety: on the role of affective states in epistemic evaluation. In U. Brun, U. Doguouglu, and D. Kuenzle (Eds.) *Epistemology and Emotions*, pp. 51–66. Aldershot: Ashgate.

James, W. (1879). Feelings of rationality. *Mind*, 4, 1–22.

James, W. (1980). *Principles of Psychology*. New York: Holt.

Kahneman, D. and Frederick, S. (2002). Representativeness revisited: attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, and D. Kahneman (Eds.) *Heuristics and Biases: The Psychology of Intuitive Judgment*, pp. 49–81. Cambridge: Cambridge University Press.

Kahneman, D. and Frederick, S. (2005). A model of heuristic judgment. In K. Holyoak and R. G. Morrison (Eds.) *The Cambridge Handbook of Thinking and Reasoning*, pp. 267–94. Cambridge: Cambridge University Press.

Koriat, A. (1995). Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General*, 124, 311–33.

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, 126, 349–70.

Koriat, A. (2000). The feeling of knowing: Some metatheoretical implications for consciousness and control. *Consciousness and Cognition*, 9, 149–71.

Koriat, A. (2006). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch, and E. Thompson (Eds.) *Cambridge Handbook of Consciousness*, pp. 289–325. New York: Cambridge University Press.

Koriat, A. and Levy-Sadot, R. (2000). Conscious and unconscious metacognition: A rejoinder. *Consciousness and Cognition* 9, 193–202.

Koriat, A. and Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 34–53.

Koriat, A., Ma'ayan, H., and Nussinson, R. (2006). The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology*, 135(1), 36–69.

Laird, J. D. (2007). *Feelings. The Perception of Self*. Oxford: Oxford University Press.

Lewis, D. (1983). Scorekeeping in a language game. In D. Lewis (1983). *Philosophical Papers*, Vol. I, pp. 233–49. Oxford: Oxford University Press.

Loussouarn, A. (2010). De la métaperception à l'agir perceptif. PhD thesis, EHESS, Institut Jean-Nicod.

Mangan, B. 2001. Sensation's ghost. The non-sensory 'fringe' of consciousness. *PSYCHE*, 7(18).

Matthen, M. (2005). *Seeing, Doing, and Knowing*. Oxford: Oxford University Press.

Metcalfe, J. (2000). Feelings and judgments of knowing: Is there a special noetic state? *Consciousness and Cognition* 9, 178–86.

Metcalfe, J. and Shimamura, A. P. (Eds.) *Metacognition: Knowing about knowing*. Cambridge, MA: MIT Press.

Nelson, T. O. (Ed.) (1992). *Metacognition. Core Readings*. Cambridge, MA: MIT Press.

Povinelli, D. (2003). Chimpanzee minds: suspiciously human? *Trends in Cognitive Science*, 7(4), 157–60.

Price, M. C. and Norman, E. (2008). *Judgment and Decision Making*, 3(1), 28–41.

Prinz, J. (2004). *Gut Reactions: A Perceptual Theory of Emotion*. New York: Oxford University Press.

Prinz, J. (2007). *The Emotional Construction of Morals*. Oxford: Oxford University Press.

Proust, J. (2006). Rationality and metacognition in non-human animals. In S. Hurley and M. Nudds (Eds.) *Rational Animals*, pp. 247–74. Oxford: Oxford University Press.

Proust, J. (2007). Metacognition and metarepresentation: Is a self-directed theory of mind a precondition for metacognition? *Synthese*, 159, 271–95.

Proust, J. (2008). Epistemic agency and metacognition: An externalist view. *Proceedings of the Aristotelian Society*, 108, 241–68.

Reber, R., Schwarz, N. and Winkielman P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, 8(4), 364–82.

Reder, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology*, 19, 90–138.

Reder, L. M. and Ritter, F. E. (1992). What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 435–51.

Reder, L. M. and Shunn, C. D. (1996). Metacognition does not imply awareness: Strategy choice is governed by implicit learning and memory. In L. M. Reder (Ed.) *Implicit memory and metacognition*, pp. 45–78. Hillsdale, NJ: Erlbaum.

Rensink, R. A. (2004). Visual sensing without seeing. *Psychological Science* 15(1), 27–32.

Rosenthal, D. (2000). Consciousness, content, and metacognitive judgments. *Consciousness and Cognition*, 9, 203–14.

Schwarz, B. L. (2002). *Tip-of-the-tongue states: Phenomenology, mechanism, and lexical retrieval*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Schwarz, N. and Clore, G. L. (1996). Feelings and phenomenal experiences. In E. T. Higgins and A. W. Kruglanski (Eds.) *Social psychology: Handbook of basic principles*, pp. 433–65. NewYork: Guilford Press.

Smith, J. D. (2005). Studies of uncertainty monitoring and metacognition in animals and humans. In H. S. Terrace and J. Metcalfe (Eds.) *The Missing Link in Cognition. Origins of Self-Reflective Consciousness*, pp. 242–71. Oxford: Oxford University Press.

Smith, J. D. (2009). The study of animal metacognition. *Trends in Cognitive Sciences*, 13(9), 389–96.

Smith, J. D., Shields, W. E., and Washburn, D. A. (2003). The comparative psychology of uncertainty monitoring and metacognition. *Behavioral and Brain Sciences*, 26, 317–73.

Spehn, M. K. and Reder, L. M. (2000). The unconscious feeling of knowing: a commentary on Koriat's paper. *Consciousness and Cognition*, 9, 187–92.

Stanovich, K. E. and West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioral and Brain Sciences*, 23, 645–726.

Tomasello, M. (1999). *The cultural origins of human cognition*. Harvard, MA: Harvard University Press.

Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28, 675–735.

Thomson, V. (2008). Dual-process theories. A metacognitive perspective. In J. St. B. T. Evans and K. Frankish (Eds.) *In two minds: Dual processes and beyond*, pp. 171–195. Oxford: Oxford University Press.

Tye, M. (2009). *Consciousness revisited*. Cambridge, MA: MIT Press.

Whittlesea, B. W. A. and Williams, L. D. (2001a). The Discrepancy-Attribution Hypothesis: I. The Heuristic Basis of Feelings of Familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(1), 3–13.

Whittlesea, B. W. A. and Williams, L. D. (2001b). The Discrepancy-Attribution Hypothesis: II. Expectation, uncertainty, surprise, and feelings of familiarity. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 27(1), 14–33.

Winkielman, P. and Fazendeiro, T. (in preparation). Effects of conceptual fluency on memory and liking.

Winman, A. and Juslin, P. (2005). 'I'm m/n confident that I'm correct': Confidence in foresight and hindsight as a sampling probability. In K. Fiedler and P. Juslin (Eds.) *Information sampling and adaptive cognition*, pp. 409–39. Cambridge: Cambridge University Press.