

Epistemic Feelings, Epistemic Emotions: Review and Introduction to the Focus Section*

Santiago Arango-Muñoz - Kourken Michaelian

Abstract: Philosophers of mind and epistemologists are increasingly making room in their theories for *epistemic emotions* (E-emotions) and, drawing on metacognition research in psychology, *epistemic* – or *noetic* or *metacognitive* – feelings (E-feelings). Since philosophers have only recently begun to draw on empirical research on E-feelings, in particular, we begin by providing a general characterization of E-feelings (section 1) and reviewing some highlights of relevant research (section 2). We then turn to philosophical work on E-feelings and E-emotions, situating the contributions to the focus section (two articles devoted to E-feelings and two devoted to E-emotions) with respect to both the existing literature and each other (section 3). We conclude by briefly describing some promising avenues for further philosophical research on E-feelings and E-emotions (section 4).

1. Introduction

1.1. E-feelings vs. E-emotions

Feelings, in general, are spontaneously-emerging occurrent phenomenal experiences, “datable states of consciousness” (Alston 1969) that form part of the stream of consciousness. Alston (1969, 5) enumerates four kinds of feelings: emotional feelings (e.g., feeling angry, anxious, or annoyed), mood feelings (e.g., feeling cheerful, gloomy, or tranquil), feelings of bodily conditions (e.g., feeling tense, sleepy, or hungry), and feelings of behavioral tendency (e.g., feeling generous, adventurous, or talkative). We can add at least three more kinds to this list: There are feelings about external states of affairs (e.g., when one feels that it is going to rain) (Dokic 2012). There are social feelings, i.e., feelings related to other people (e.g., when one feels that one can trust someone, or when one feels that someone is lying) (de Sousa 2008; Dokic 2012). Finally, there are *epistemic* feelings (de Sousa 2008; Dokic 2012) – feelings concerning the subject’s own mental capacities and mental processes.

* Thanks to Carla Bagnoli for her guidance on the preparation of the focus section and to our contributors for respecting a tight timeline.

Epistemic *feelings* and epistemic *emotions* are often not clearly differentiated in the literature, with researchers using the terms more or less interchangeably. There are, however, some key differences between feelings and emotions. First, whereas feelings are phenomenal experiences of which the subject is conscious (though they often form part of the fringe of consciousness – see section 2.1), emotions can in many cases be unconscious – they need not form part of the stream of consciousness. Second, emotions, unlike feelings, in many cases involve sophisticated cognitive states (including beliefs and desires) – unlike feelings, emotions often involve properly propositional content (see section 1.4 on the content of E-feelings). Given the existence of these basic differences between feelings and emotions in general, there is a strong *prima facie* case to be made for distinguishing between E-feelings and E-emotions.

Though E-feelings and E-emotions, as affective phenomena playing important motivational and regulative roles in our cognitive lives, have enough in common to justify discussing them together, the prudent strategy, given the deep differences between them and absent a unifying account, is to distinguish clearly between them. Since Meylan's contribution to the focus section provides a systematic discussion of the key features of E-emotions, and since an adequate philosophical approach to E-feelings must cohere with the available psychological research in the area, research with which many philosophers will be unacquainted, we begin by providing some background on E-feelings, coming back to E-emotions in section 3.

1.2. Four (or five) examples

We begin by discussing some specific E-feelings: the feeling of confidence, the feeling of knowing (and the related tip-of-the-tongue state), the feeling of error, and – an emerging area of investigation – the feeling of forgetting.

1.2.1. The feeling of confidence

While returning from a holiday in Madrid, the first author recently had a confusing experience. As he arrived at the airport, he discovered that he didn't have his passport with him. Since he was sure that he had verified that he had packed it the night before, this left him quite puzzled. Fortunately, he was able to complete his journey home without the passport. But where was the passport? Though he felt confident that he had seen it the night before, all the available evidence pointed to the fact that his memory was incorrect. Two weeks later, however, he received a call from Madrid: his passport had been found behind a table in his hotel room. At that point, things fell into

place: His feeling of confidence in his memory was justified – he had in fact seen his passport the night before his departure. But he now remembered that, rather than packing it, he had placed the passport on the table next to his hand luggage, so that he would see it when he woke up. He must, without realizing it, have knocked the passport off of the table in his rush to leave on time.

The relationship between episodic memory and confidence has traditionally been studied in relation to eyewitness reports (see Krug 2007 for a review). In this context, researchers have often failed to find a positive correlation between correct answers and confidence, i.e., subjects tend to be overconfident about their memory performance. However, other studies dealing with semantic memory (e.g., Perfect 2004; Luna and Martín-Luengo 2012) and visual perception (see Fleming and Dolan 2012 for a review and discussion; see also Yeung and Summerfield 2012) have found positive correlations.

1.2.2. The feeling of knowing and the tip-of-the-tongue state

Imagine that you're participating in a TV quiz show such as *Question pour un champion* or *Jeopardy*, where participants compete to answer general knowledge questions. Each participant should press the buzzer if and only if he knows the answer to the given question: pressing the buzzer gives one the opportunity to earn points by answering the question, but if one presses the buzzer but is then unable to answer the question, one loses points. Moreover, each participant, if he is going to press the buzzer, should do so as quickly as possible, since if another contestant presses it first, he loses the opportunity to answer the question. In this situation, you press the buzzer, in some cases, because the answer to the question immediately comes to mind. But in many cases you press the buzzer before having retrieved the answer, relying on a gut feeling that tells you that you'll easily be able to retrieve it. This is the feeling of knowing (FOK).¹

Closely related to the FOK is the tip-of-the-tongue phenomenon (TOT), which each of us knows from his own experience. Consider any case in which you feel that you know the answer to a question – for example, someone's name or phone number – despite having just failed to retrieve it. Notwithstanding the failure to retrieve the information, you feel as if the information is available in memory and possibly that you're likely to be able to retrieve it if given enough time. This is the tip-of-the-tongue state.

¹ Note that the feeling of knowing is a feeling concerning the possibility of retrieving information from memory, independent of whether the information in question is true or false, justified or unjustified. Hence the FOK is not, strictly speaking, a feeling that one *knows*, in the epistemologist's sense.

There is some ambiguity in the way the FOK and the TOT are defined in the literature, with many researchers using FOK to refer to the feeling that a subject has when, after failing to retrieve an item, he nevertheless feels that it is available in memory (e.g., Hart 1965; Nelson and Narens 1990; Nelson 1999). The problem with this approach is that it makes the FOK equivalent to the TOT, thus leading researchers to investigate both feelings by asking the subject to rate the probability that she will recognize the target word among distractors, after retrieval failure (Spehn and Reder 2000). Reder's (1987; 1988; 1996) definition of the FOK as an experience the subject undergoes *before* attempting to recall information avoids this problem; the TOT is then characterized by its occurrence after a memory retrieval failure together with the feeling of immediacy of the answer (Schwartz and Metcalfe 2011). Though it is likely that the TOT is preceded and partly constituted by the FOK (Koriat and Levy-Sadot 2001; Mangan 2000; Moulin and Souchay 20xx), this approach allows us to differentiate them functionally.

1.2.3. The feeling of error

Consider the following feelings:

- The feeling that one sometimes gets, after solving a math problem, that one has miscalculated.
- The negative feeling that sometimes arises after a reasoning process, warning one of possible incoherence in one's reasoning.
- The feeling that there is something wrong with a perceptual experience, warning one of a potential illusion or hallucination.

These are all cases of the feeling of error – an apparently groundless feeling that emerges into consciousness, independently of the available evidence, and that points towards an error in one's mental processes. The feeling of error can be defined as the subjective experience that something went wrong during the execution of a mental action (e.g., reasoning or decision making; see the discussion of post-evaluation in section 3.1.2).

Recent research on executive control has investigated errors in bodily action under the heading of error monitoring; see Wessel 2012 for a review. For feelings of error following reasoning tasks, see De Neys *et al.* 2011; De Neys 2012.

1.2.4. The feeling of forgetting

A final case: A friend travelling to the south of France for a cycling tour experienced a strange feeling as he left the train station and got into the car that would take him to the starting point of the tour. He felt weird, uncomfortable, and even slightly sad, but he was initially unable to identify the cause of the feeling. As the feeling gradually became clearer, he realized that he felt like he had forgotten something, though he didn't yet know what. Only too late, as he was getting out of the car at the starting point of the tour, did he realize that he had forgotten his favourite hat on the train.

The feeling of forgetting is a new but promising area of research; see Halamish *et al.* 2011 for an empirical approach and Arango-Muñoz (2013b) for theoretical discussion.

1.3. Other E-feelings

Many additional E-feelings have been investigated (or at least proposed), including the following.

THE FEELING OF UNCERTAINTY: The subjective awareness of the imprecision of a mental representation or piece of information (Smith 2009; Bach and Dolan 2012; and see Dokic, this issue).

THE FEELING OF FAMILIARITY: The sense of “having prior experience, whether or not one actually has it” (Whittlesea 1993; Whittlesea and Williams 2000).

THE FEELING OF UNDERSTANDING: A feeling of intellectual satisfaction that motivates the endorsement of an explanation, a sense that we have achieved an understanding of a phenomenon that was not clearly understood before (Gopnik 1998, 2000; Trout 2002, 2007). This is sometimes called the “ah ha” feeling (Mangan 2001) or the eureka feeling.

THE FEELING OF DIFFICULTY: The subjective perception of the difficulty of a task (Efklides 2002; Efklides and Touroutoglou 2010).

THE DÉJÀ VU EXPERIENCE: The subjective impression of the familiarity of a present visual experience, relative to an undefined past experience (Brown 2003; Kusumi 2006).

THE BLANK IN THE MIND EXPERIENCE: A feeling concerning prospective memory (memory for tasks to be performed in the future), this is “a sudden awareness of having no content in conscious awareness ...the person feels that s/he has lost

track of the intention or the cue for initiating one's [thought or] action" (Efklides and Touroutoglou 2010).

THE FEELING OF CURIOSITY: An emotional state that motivates exploratory behaviours and knowledge acquisition (Litman *et al.* 2005) (but see Inan 2012 for a different view).

THE FEELING OF RIGHTNESS: The subjective experience that the execution of a mental action (e.g., reasoning or decision making) was successful (Mangan 1993, 2001; Thompson 2009; Thompson *et al.* 2011).

THE FEELINGS OF COMPETENCE: The feeling that one is able to carry out a given mental action (Bjork 1999).

THE FEELING OF PRESENCE: The inarticulate feeling that subjects have when they are visually conscious of three-dimensional objects: "The feeling of presence is similar to assertion: attached to a visual scene, the feeling of presence asserts it, so to speak – it makes one feel that the scene being described is present" (Matthen 2005: 305).

THE FEELING/SENSE OF AGENCY OVER THOUGHTS: The impression that one's own mental states are caused by oneself. This feeling may also points towards the control that we have of our thoughts, and a lack of it may produce mental pathologies such as thought insertion (Proust 2006, 2009a; Carruthers 2012).

THE FEELING OF RATIONAL RELATION OR RATIONALITY: James introduced this concept to designate the sense of adequacy of a sentence, argument, or linguistic construction (James 1890).

EPISTEMIC ANXIETY: "A force that normally determines how much evidence we are inclined to collect and how thoroughly we will weigh it before making up our minds" (Nagel 2010).

THE FEELING OF IMMANENCE: In perceptual processes, the feeling that much more detailed information is available on the periphery for retrieval if needed (Mangan 2001).

THE SENSE OF MEANING OR SEMANTIC EXPERIENCE: "When (for example) one hears someone speak in a language one understands" (Strawson 2011) (cf. James 1890; Mangan 1993, 2000, 2001). "Something happens in your mind – some act of consciousness – over and above the hearing of the words, some act of consciousness which may be called understanding their meaning" (Moore 1962: 281).

THE FEELING OF PASTNESS: The feeling that a representation is a representation of a past experience, that is, memory. According to Russell, “[t]here may be a specific feeling which could be called the feeling of “pastness”, especially where immediate memory is concerned” (Russell 1921: 162).

This list does not aim to be exhaustive, nor do we claim that each of the items on the list is a genuine epistemic feeling; but hopefully the list will give the reader a sense of the potential scope of the category.

1.4. Content and function

What do the various E-feelings have in common with each other? In this section, we sketch a partial and tentative answer to this question, focussing on the content and function of E-feelings.

CONTENT: While E-feelings are normally caused by contentless cues and heuristics, the feelings themselves often condense implicit knowledge or information (Norman *et al.* 2010). Taking Proust’s distinction between metacognition and mindreading into account, the content of E-feelings should be characterized in nonconceptual and non-metarepresentational terms (Proust 2007, 2012, 20xx). We have already said (section 1.1) that part of this content concerns the subject’s own mental capacities and processes. The additional key ingredient is positive or negative *affect* directed at these processes (Proust 2009b). The affective component of the E-feeling evaluates the mental process to which it is directed and indicates how well it will unfold, is unfolding, or has unfolded (see section 3.1.2). As Dokic puts it, the content can be described as evaluating the competence of an agent in the execution of a cognitive action: “I can do this” or “this can be done” (Dokic 2012).

FUNCTION: To survive and fulfil their needs in an uncertain world, agents need a means to learn about the structure of the world, to predict what is going to happen, and to react adequately. Cognition is (among other things) a means of coping with the uncertainty of the world. Analogously, we can define metacognition as a means to cope with the uncertainty of the mind (Proust 2007, 2008, 20xx). Subjects do not have perfect self-knowledge (knowledge of their own minds) and therefore they are often uncertain about their own mental states and mental capacities; against this background, feeling-based metacognition evolved to help the subject to cope with mental uncertainty.

There are at least two varieties of mental uncertainty in play here, which we can illustrate using the example of memory. First, subjects seem to have a general sense of what they are able or unable to remember. At the same time, they do not, prior to retrieval, have direct access to the contents of

their memories (see section 2.3.1). How, then, do subjects know whether they will be able or unable to remember a given item? Michaelian (2012a) refers to the problem of determining whether one should rely on one's memory or, instead, on some alternative resource as the *selection problem*. Arango-Muñoz has recently argued that subjects may solve the selection problem primarily by relying on E-feelings (Arango-Muñoz 2013b). On this view, E-feelings – for example, the FOK – are the solution that evolution has designed to allow subjects to cope with uncertainty about the contents of their own memories.

Second, in addition to uncertainty about whether one will be able to retrieve the needed information, subjects must cope with another kind of uncertainty after having retrieved an item from memory. Given that retrieval may output inaccurate information (due both to the fact that memory stores inaccurate information and to the reconstructive character of the retrieval process itself), the subject must decide whether to accept retrieved information. Michaelian referred to this as the *endorsement problem* (2012a; 2012b). Though he suggested that subjects solve the endorsement problem by relying on the sort of non-feeling based metacognitive monitoring described by the source monitoring framework (Johnson *et al.* 1993; Mitchell and Johnson 2009), it is likely that E-feelings also play an important role here (Arango-Muñoz 2013b) (see section 4 for further discussion). For example, the feeling of rightness and the feeling of error provide quick assessments of retrieved information.

2. *Psychological research on E-feelings*

In this section, we review psychological research on E-feelings, beginning with their role in the fringe of consciousness, covering foundational work in the 1960s-1980s, and then looking at recent trends.

2.1. The fringe of consciousness

Though E-feelings are sometimes phenomenally salient (for example, when a subject experiences a TOT state while struggling to recall information from memory), they typically have a faint and fleeting character, and the subject does not attend closely to them. Thus recently a number of researchers (Mangan 1993, 2000, 2001; Norman *et al.* 2010; Dokic 2012) have suggested viewing E-feelings in terms of the concept of the “fringe of consciousness”, which James introduced to designate “the influence of a faint brain-process upon our thought, as it makes it aware of relations and objects but dimly perceived” (James 1890, vol. 1: 258). In contrast to the nucleus of consciousness, the fringe is composed of contents that are not currently the focus of attention.

Mangan (1993; 2000; 2001) has developed a thorough functional analysis of the fringe of consciousness. While we do not have space to review Mangan's analysis in detail here, we note that he explains some of the phenomenal characteristics of E-feelings – in particular, their fleeting character and their low resolution – in terms of cognitive economy. The idea is that E-feelings have their particular character due to the limited capacity of working memory, which according to many theorists is the locus of the stream of consciousness (Baars 1997; Dehaene 2001; Dennett 2005). Mangan's analysis also suggests that E-feelings are intentional, condensing information that guides subjects' behaviour. (Norman and her colleagues refer to this as the "summary function" (Norman *et al.* 2010).) Finally, in line with our discussion above (section 1.4), Mangan emphasizes that E-feelings are evaluative: they reveal to the subject her relation to a mental goal, such as retrieving a word or solving a mental problem. By means of E-feelings, the subject becomes aware of the success or failure of her mental activities in attaining a mental goal. As James put it, "[t]he most important element of these fringes is ...the mere feeling of harmony or discord, of a right or wrong direction in the thought" (1890, vol. 1: 261).

2.2. 1960s-1980s

Although James was already moving towards the notion of E-feelings at the end of the 19th century, it wasn't until the latter half of the 20th century (after the behaviourist interlude) that the experimental investigation of E-feelings really got underway. In this section and the next, we review selected highlights of this experimental work.

2.2.1. The direct access model

Hart (1965) was the first to study the FOK in an empirical setting, looking at the relation between the FOK and memory performance by asking subjects to make judgements about the future memorability of currently unavailable items. He found that the likelihood of correctly recognizing a nonrecalled item among distractors was higher for items that elicited a FOK than for items that did not. Hart explained these results by postulating the existence of an internal mechanism that has direct access to the contents of memory. According to this model, although the subject himself does not have access to the contents, the metacognitive mechanism does, and the FOK is elicited based on this access.

2.2.2. The tip-of-the-tongue phenomenon

Around the same time, Brown and McNeill published their seminal paper on the TOT phenomenon (Brown and McNeill 1966). They investigated the TOT by reading definitions of low frequency words to subjects and asking them to recall the corresponding words. If the subject reported a TOT, the experimenter asked further questions: How many syllables does the target word have? What is the first letter? What other words sound like the intended word? What other words have a similar meaning? It turned out that, when subjects reported a TOT, they also had partial or generic knowledge of the missing word; e.g., they often knew the first letter of the missing word, the number of syllables in it, or the primary stress pattern; in addition, most of the reported phonologically similar words actually matched the target word.

2.2.3. A developmental perspective

Flavell's influential work on metacognition dealt mainly with the development of metacognitive understanding in children, that is, with the understanding that children acquire about their own and others' minds during the first years of life. Although he did not work directly on metacognitive experiences, he did suggest that feelings and experiences played an essential role in the cognitive economy and modulated metacognitive control: "Many metacognitive experiences have to do with where you are in an enterprise and what sort of progress you are making or are likely to make: You believe/feel that you have almost memorized those instructions, are not adequately communicating how you feel to your friend, are suddenly stymied in your attempt to understand something you are reading, have just begun to solve what you sense will be an easy problem, and so forth" (Flavell 1979: 908). Despite the fact that Flavell hinted that feelings and emotions could play an important role in metacognition, however, did not develop a theoretical account of E-feelings or experience-based metacognition.

2.2.4. Towards a functional framework

Nelson and Narens (1990) were the first to propose a systematic functional account of metamemory. In their framework, cognitive processes occur either at the object level or at the meta-level. The object level includes core cognitive processes, such as perception, reasoning, and memory. The meta-level has two key elements: (1) the particular way a subject becomes aware of what happens in her mind (monitoring), and (2) the subject's actions and reactions caused by such awareness (control). What this model suggests, in other words,

is that metacognitive feelings are the outputs of a monitoring mechanism, and modulate control behaviors. We may take Nelson and Narens' model to apply mainly to what Arango-Muñoz calls "low-level metacognition" (2011), that is, the kind of metacognition based on E-feelings rather than second-order thoughts or metarepresentations (cf. Koriat 2000 on experience-based vs. information-based metacognition).

2.3. 1990s-present

The past two decades were an especially fruitful period for the research on metacognition and metacognitive feelings; we can review only selected themes here.

2.3.1. The inferential model

Common to most researchers in this period is the rejection of Hart's direct access model and the endorsement of new models based on inference and heuristics, the core idea being that the monitoring mechanism relies on internal and external cues to infer information about the subject's mental states and processes.² Proposed cues include the accessibility of related information (Koriat 1993), fluency of processing (Whittlesea and Williams 1998, 2001), and familiarity (Reder and Ritter 1992; Reder 1996; Metcalfe 1993). For example, if the subject is unable to respond to a question but the question is familiar, the monitoring mechanism may infer that the subject knows the answer, producing a TOT state. This turn was motivated, *inter alia*, by the emergence of a view of remembering as a constructive process, in which retrieval effectively generates new representations, rather than simply accessing existing stored information (Schacter and Addis 2007; Michaelian 2011, 2013).

2.3.2. Metaperception and metareasoning

Metacognition research has tended to focus primarily on metamemory, but there is increasing interest in other forms of metacognition. For example, Levin initiated research on visual metacognition or metaperception (2004; see

² We emphasize that the inferences in question have little in common with inference as typically viewed by philosophers. While philosophers tend to regard inference as a conscious logical operation over propositions in which the subject reaches a conclusion on the basis of explicit premises or evidence, the inferences involved in metacognition are unconscious operations of the mind/brain, which need not be logical in nature. To illustrate, consider the way in which the visual system infers a third dimension – depth – from the bidimensional information received via the retina, an inference which is neither conscious nor governed by the rules of logic.

also (Loussouarn *et al.* 2011; Dokic and Martin 2012). Thompson (Thompson 2009; Thompson *et al.* 2011) and De Neys (De Neys *et al.* 2011; De Neys 2012) have adapted classic reasoning tasks to test subjects' metareasoning abilities and their subjective awareness of the correctness of their reasoning. In the domain of problem solving, Efklides and her colleagues have developed paradigms to test the feeling of difficulty, the blank in the mind phenomenon, and related feelings (Efklides 2002; Efklides and Touroutoglou 2010).

2.3.3. New perspectives on metamemory

A promising new development in metamemory research is the embodied approach to E-feelings. This bottom-up perspective explores the way somatic cues such as facial gestures or subliminal buzzes shape phenomenal E-feelings and judgements (Goldinger and Hansen 2005; Koriat and Nussinson 2009). Koriat and Nussinson (2009), for example, found that asking subjects to contract the corrugator muscle (by frowning) while carrying out a cognitive task caused the experience of mental effort. They found that FOK co-varies with the distention of the corrugator muscle, whereas the feeling of difficulty or uncertainty co-varies with its tension.

At the same time, a complementary top-down approach has investigated the effects of the way metacognitive questions are framed (Koriat *et al.* 2004). Finn (2008), for example, demonstrated that asking subjects questions in terms of forgetting instead of remembering reduces the oft-observed overconfidence effect. Framing effects seem to play a more robust role in metacognition and E-feelings than was previously thought: the way you think about your mind and the concepts you use to think about it influence the way you feel about it.

3. *E-feelings and E-emotions in philosophy*

We turn now from psychology to philosophy, looking at philosophical work on E-feelings and E-emotions.

3.1. Philosophical work on E-feelings

In contrast to the level of activity in psychology, relatively little has been written about E-feelings in philosophy. However, there have been a number of important developments in recent years. (We can here address only selected contributions; in particular, space does not permit us to discuss Rosenthal's seminal work (2000; 2012) on consciousness and the FOK/TOT.)

3.1.1. An embodied account

Dokic (2012; see also his contribution to this issue) has proposed an embodied account of E-feelings, the “water diviner” model. On his view, E-feelings are first and foremost bodily experiences, i.e., experiences about bodily states. They are diffuse affective states registering internal physiological conditions and events. But just as the water diviner’s sensations reliably co-vary with physical conditions, namely the presence of underground water, E-feelings reliably co-vary with mental conditions. For example, the FOK – which is, in this view, essentially a bodily feeling – reliably co-varies with the fact that a given piece of information is stored in the subject’s memory. This reliable co-variation explains why self-ascription of mental states based on bodily experiences can lead to self-knowledge.

3.1.2. Mental action

Proust (Proust 2001, 2006, 2007, 2008, 2009a, 20xx) has investigated the relation of E-feelings to mental agency. On her view, the capacity to perform a mental action is partially constituted by metacognitive evaluations of (1) the possibility of carrying out the mental action (“self-probing” or “prospective monitoring”) and (2) the adequacy of the action once performed (“post-evaluation” or “retrospective monitoring”).³ Such evaluations are conveyed by E-feelings. Not only do E-feelings convey these evaluations, but they are also affective markers that provide a sense of agency over thought (i.e., the subjective sense that the agent herself is the cause of mental events) (Proust 2006, 2009a).

3.2. Philosophical work on E-emotions

In addition to developing theories of the nature and role of E-feelings, a number of philosophers have looked at E-emotions.

3.2.1. The epistemic role of feelings and emotions

De Sousa (2008; 2011) has argued that feelings and emotions are key to solving problems including the frame problem and the problem of epistemic justification. On his view, feelings and emotions are patterns of saliency among objects of attention and strategies, and in this way they close the computational and epistemic gaps. They are the product of a subpersonal intuitive system that provides premises to the analytical system that then makes use of

³ Arguably, we should add a third stage of metacognitive evaluation, occurring during the execution of a mental action and monitoring its unfolding (Arango-Muñoz 2013a).

them in explicit inferences. According to de Sousa, familiar emotions such as fear, greed or trust can also function as E-feelings by affecting conviction, inference, or the way we use cognitive strategies (cf. Hookway 2008, 2003; Nagel 2010). In contrast to our approach here (and Meylan's in her contribution to the focus section), de Sousa's approach blurs the distinction between E-feelings and E-emotions.

3.2.2. Emotions and virtues

Focussing on the connection between emotions and virtues, Morton (2010; 2013) has arrived at broadly compatible conclusions about the role of E-emotions in shaping belief-acquisition. He contends that, while E-emotions are connected to epistemic virtues, they have an independent role to play: one can, for example, be an excellent researcher, but if one lacks genuine curiosity about one's research area, one will still be limited in certain respects, including willingness to take intellectual risks that might either drive the field forward or lead nowhere. In general, he argues, while epistemic virtues might in principle operate without the corresponding epistemic emotions, there are many circumstances in which cognizers with our basic intellectual constitutions need to be driven by emotion if they are to sustain the operation of epistemic virtue.

3.3. The focus section

The contributions to the focus section advance these recent lines of philosophical inquiry into the nature and role of E-feelings and E-emotions.

3.3.1. Safety and margins for error

Building on his embodied account (Dokic 2012), Dokic's contribution explores the epistemic role of E-feelings. He begins by situating epistemic feelings with respect to the dual system/dual process framework (Frankish 2010; Evans and Stanovich 2013), according to which subjects by default employ type 1 processes (heuristic, unconscious, automatic), only sometimes switching to type 2 processes (systematic, conscious, deliberate): E-feelings are cross-level states, produced by implicit, type 1 monitoring but available to participate in explicit, type 2 reasoning (cf. Koriat 2006).

Drawing on Proust (2007), Dokic emphasizes that, while the content of E-feelings may refer to first-order representations, it is not strictly metarepresentational: it does not represent representations as representations, for the implicit monitoring mechanisms which produce E-feelings are not sensitive to the content of the processes they monitor but only to properties such as fluency

(Oppenheimer 2008). The fact that metacognitive mechanisms are insensitive to the content of the processes they monitor gives rise to a puzzle about the epistemic role of the feelings they produce: how, for example, can the FOK play an epistemic role, given that it is produced without direct access to the relevant stored information?

Focussing on the case of perceptual uncertainty monitoring (Smith *et al.* 2003), Dokic proposes an account of the epistemic role of E-feelings in terms of the notions of safety and margins for error. Many epistemologists have argued that safety is required for knowledge – roughly, that a true belief amounts to knowledge only if it was produced by a mechanism that would not easily have produced a false belief under the relevant circumstances (Sainsbury 1995). The notion of safety, in turn, is related to that of a margin for error (Dokic and Égré 2009); for example, in the case of visual perception, if the subject correctly judges that a given stimulus in a series has a certain property, his visual belief will be safe only if he does not tend to incorrectly judge that the next stimulus in the series has it if it does not – that is, if he has a margin for error. Dokic's core claim is that, in the case of perception, the feeling of certainty tracks the safety of perceptual beliefs: if the subject feels certain about her classification of a stimulus, the corresponding belief is normally safe, and the subject has a margin for error. Similarly, the feeling of uncertainty tracks lack of safety and absence of a margin for error.

3.3.2. Mental action and self-ascription

Like the work by Proust described above (section 3.1.2), Arango-Muñoz's contribution explores the relationship between E-feelings and mental action. He points out that mental action and self-ascription are distinct, separable consequences of E-feelings, and aims to describe the relationship among these three factors. He begins by reviewing Dokic's argument (Dokic 2012) for the insufficiency of the "ascent routine" to account for cases of non-transparent self-ascription, in which the subject ascribes a mental state to himself without having access to the content of that state. Employing the ascent routine, one is able to determine whether one believes that *P* simply by considering the same evidence that one would consider to determine whether *P* (Evans 1982; Moran 2001). As Dokic shows, in cases other than yes-no questions, the subject can be in a position to ascribe a mental state to himself without having access to the content of that state.

Rather than relying on the ascent routine, non-transparent self-ascription depends on E-feelings; for example, one can ascribe knowledge to oneself by relying on one's FOK without having access to the relevant stored information.

While Arango-Muñoz accepts Dokic's basic strategy, he argues that it is incomplete: in light of the embodied account of E-feelings, how, exactly, does the subject move from a given E-feeling to a given self-ascription? In order to answer this question, he draws on his two-level model of metacognition (Arango-Muñoz 2011). In *low-level* metacognition, E-feelings provide an implicit, affective assessment of a given cognitive task involving a non-transparent object. Since the object is non-transparent, the subject is in the same position with respect to his own mind that he is in with respect to the minds of others; thus, in *high-level metacognition*, he turns his mindreading mechanism upon the E-feelings produced by low-level metacognition to self-ascribe a mental state.

Invoking the mindreading mechanism, however, does not yet provide a full account of how subjects move from E-feelings to self-ascriptions. Arango-Muñoz therefore turns to the (learned) epistemic rules (Byrne 2005) governing the relations among E-feelings, self-ascriptions, and mental actions. On his account, it is such rules (possibly implicit, rather than explicitly followed by the subject) that guide the subject's response to his E-feelings. On the one hand, there are *epistemic rules for action* that determine what mental action one should perform given a certain E-feeling. On the other hand, there are *epistemic rules for self-ascription* that determine what to believe about one's mental states given a certain E-feeling. For example, there may be an epistemic rule for action guiding the subject to attempt to retrieve the needed information if he has an FOK; at the same time, there may be an epistemic rule for self-ascription guiding the subject to form the belief that he can remember if he has an FOK. Of course, as Arango-Muñoz grants, the interaction between rules for action and rules for self-ascription may be complex, and this may require further investigation

3.3.3. Feelings vs. emotions

In her contribution, Meylan aims to delineate the category of epistemic emotions more precisely than has been done so far, paying particular attention to differences between E-emotions and E-feelings. She points out that philosophers and psychologists have often grouped the sorts of E-feelings discussed above (section 1) – the FOK, the feeling of familiarity, the feeling of certainty, and so on – with phenomena such as curiosity (Loewenstein 1994), interest (Silvia 2006), surprise (Lorini and Castelfranchi 2007), and trust (Origi 2008). She argues that E-feelings should be distinguished from E-emotions, and goes on to explore whether curiosity, interest, surprise, and trust form a coherent kind.

Meylan argues that emotions in general are characterized by five essential features. First, they have a felt character – that is, episodes of emotion are felt by the subject. Second, they involve characteristic physiological changes, for example, changes in facial expression. Third, they have either a positive or a negative valence (as in the cases of happiness and anger, respectively). Fourth, they have specific intentionality: a given episode of emotion presents a definite object (e.g., a situation) to the subject as having a definite evaluative property (e.g., being dangerous). Finally, episodes of emotion are subject to two standards of evaluation: they can be correct or incorrect (depending on whether the object of the episode in fact has the evaluative property the episode presents it as having); and they can be justified or unjustified.

On the basis of this characterization of emotions, Meylan argues that E-feelings are not emotions. We note that while, for reasons given in section 1, we agree with this claim, we disagree with Meylan's reason for the claim: she argues that E-feelings do not present subjects with evaluative properties, but our view of E-feelings sees them precisely as presenting the subject with evaluations (see also Dokic 2012; Proust 2009b; Arango-Muñoz 2013a, as well as Dokic's and Arango-Muñoz's articles in this issue). She likewise argues that curiosity may not be a genuine emotion, on the ground, first, that it need not present the subject with an evaluative property and, second, that, it can be satisfied or unsatisfied, which makes it look more like a desire than an emotion. Of the phenomena considered by Meylan, only interest, surprise, and (felt) trust appear to have all five of the properties she takes to be necessary features of emotions.

Finally, Meylan offers another reason for not grouping the candidate E-emotions she surveys together as members of a natural kind: regardless of whether they are all *emotions*, they do not all appear to be *epistemic*. Her argument here appeals to the "formal object standard of epistemicity", the idea that an emotional episode is epistemic if and only if its formal object is an *epistemic* evaluative property. If we assume that an evaluative property counts as epistemic if it bears a specific relation to truth, Meylan argues, it appears unlikely that that interest, surprise, and trust constitute a natural kind, given their different relations to truth (for example, something need not be true in order to be interesting). Thus the article ends on a pessimistic note: not only should E-feelings and E-emotions be distinguished, but even what one might take to be relatively uncontroversial cases of E-emotions may not constitute a coherent domain of investigation.

3.3.4. Emotions and virtues: A social perspective

Unlike Meylan, Morton takes the epistemicity of an emotion to be determined by its connection to belief formation, rather than directly to truth. Building on his earlier work on the relation between emotions and virtues (Morton 2010, 2013), he sketches an optimistic picture of the role of emotions in inquiry, focussing on the way in which apparently epistemically unworthy or vicious emotions – such as nosiness, obsessiveness, wishful thinking, denial, and partisanship – if they are appropriately distributed, can turn out to have beneficial epistemic effects.

Morton argues that, just as emotions in general are intrinsically neither virtuous nor vicious, any given E-emotion can motivate either epistemic virtue or epistemic vice. Despite this neutrality, he maintains, E-emotions as a matter of fact tend naturally to give rise to vices; thus the fact that we have learned how to harness our emotions so that they play a productive role in inquiry represents a significant accomplishment. But we have indeed accomplished this, according to Morton: inquiry may actually go better if agents are motivated by unworthy emotions than if they are motivated by a passion for impersonal objectivity.

Morton begins by pointing out that human social life involves constant attacks on and defences of epistemic authority, and argues that E-emotions – emotions of partisanship, self-assertion, and epistemic denigration – play a crucial role in maintaining epistemic authority. The role of these emotions, in turn, gives rise to a tension between accuracy and authority, and the need to protect one's authority can then lead to the vice of epistemic self-indulgence, in which one works on isolated projects at the expense of collective projects, as a way of protecting one's authority. Drawing on a series of examples, Morton argues, however, that the same vice can, where there are competing groups of inquirers, aid in the attainment of objectivity (cf. Goldman 2002).

4. *Avenues for future research*

The contributions to the focus section suggest a number of promising avenues for future research on E-feelings and E-emotions.

WHAT MAKES FEELINGS/EMOTIONS EPISTEMIC? In this introduction, we defined E-feelings as feelings concerning the subject's own mental capacities and processes. Meylan defines E-emotions as emotions concerned with evaluative properties tightly linked to the goal of truth. Morton defines E-emotions as emotions concerned with belief-formation

more broadly. Which of these non-equivalent criteria for the epistemicity of feelings/emotions is preferable remains to be determined.

WHAT IS THE RELATIONSHIP BETWEEN E-FEELINGS AND E-EMOTIONS? While, given the differences between E-feelings and E-emotions canvassed in this introduction and in Meylan's article, it seems unlikely that E-feelings and E-emotions constitute a unified category, as affective phenomena that play important roles in regulating cognition, they nevertheless clearly overlap to some extent, and future work on the nature of the relationship between them would be welcome.

WHAT ARE THE RELATIONSHIPS AMONG E-FEELINGS? WHAT ARE THE RELATIONSHIPS AMONG E-EMOTIONS? Similarly, there is room for additional work on relationships among (putative) E-emotions, of the sort done by Meylan in her contribution, and likewise for work on relationships among E-feelings.

HOW ARE E-FEELINGS AND E-EMOTIONS RELATED TO CORE EPISTEMOLOGICAL CONCEPTS? Morton has done considerable work to isolate the role of E-emotions both at the level of individual epistemology and at the level of social epistemology. Similarly, Dokic and Arango-Muñoz have begun to explore the role of E-feelings in knowledge-acquisition. But there remains, in our view, considerable work to be done in this vicinity. For instance, Morton's claim that individual E-emotions are not intrinsically connected to specific cognitive virtues or vices may be open to challenge, and additional empirical evidence in favour of Dokic's and Arango-Muñoz's claims about the reliability of E-feelings should be sought.

WHICH OF THE PUTATIVE E-FEELINGS LISTED IN SECTION 1.3 ARE GENUINE? Above, we listed a number of potential E-feelings. Some of the relevant phenomena, however, might best be grouped separately from E-feelings. For example, in work in progress, Michaelian treats the feeling of pastness (Russell 1921) as an epistemic feeling; on his view, however, the monitoring mechanism responsible for producing the feeling of pastness is not insensitive to the content of the memory processes that it monitors, marking it off from other E-feelings, which are normally caused by contentless cues.

DOES LOW-LEVEL METACOGNITION NECESSARILY RELY ON E-FEELINGS? In this introduction, we have tended, in line with Dokic's and Arango-Muñoz's views, to treat low-level (unconscious, automatic) metacognition as always involving E-feelings (Arango-Muñoz 2011). However, this is not a conceptual necessity, and there are approaches to low-level metacognition which characterize it in terms of the operation of heuristic monitoring which does not output E-feelings (Johnson *et al.* 1993; Michaelian 2012a). Is there simply a gap in such accounts, or does low-level metacognition in fact sometimes operate without producing E-feelings?

HOW CAN RIVAL EPISTEMOLOGICAL FRAMEWORKS ACCOMMODATE E-FEELINGS AND E-EMOTIONS? As the four contributions to the focus section make clear, E-feelings and E-emotions play a number of important roles in our epistemic lives, but the ability of rival epistemological frameworks to accommodate these roles remains largely to be explored. Of particular interest here is the stance that epistemological internalists should take towards feelings and emotions. Can foundationalism, for example, allow epistemic feelings to play something like the role sometimes assigned to perceptual seemings? This remains an open question (but see Proust 2008), as most epistemological work on E-feelings and E-emotions so far is broadly externalist in spirit (for example, Dokic relates E-feelings to safety, while Morton looks at the relationship of E-emotions to virtue).

HOW SHOULD THE CONTENT OF E-FEELINGS BE CHARACTERIZED? Meylan argues that E-feelings do not present the subject with evaluative properties. On Dokic's and Arango-Muñoz's accounts (see also Proust 2008, 2009b), however, they do precisely that, indicating the subject's capacity to successfully perform some cognitive task or evaluating the success of a cognitive process. While it isn't clear at this stage how to resolve the disagreement, the existence of the disagreement does point to the necessity for further work on the content of E-feelings.

CAN DOKIC'S ACCOUNT BE EXTENDED BEYOND PERCEPTUAL BELIEFS? Dokic argues that the feeling of certainty/uncertainty tracks the safety/unsafety of perceptual beliefs. He also briefly argues that conceptual (as opposed to perceptual) certainty/uncertainty may play a similar role, tracking the

safety/unsafety of the conclusion of an inference. Can this suggestion about the role of E-feelings in inference be developed more fully? Can the account be extended to cover other epistemic sources (e.g., memory and testimony)?

References

- Alston, William P., 1969, "Feelings", in *The Philosophical Review*, 78, 1: 3-34.
- Arango-Muñoz, Santiago, 2011, "Two levels of metacognition", in *Philosophia*, 39, 1: 71-82.
- Arango-Muñoz, Santiago, 2013a, (forthcoming), "The nature of epistemic feelings", in *Philosophical Psychology*.
- Arango-Muñoz, Santiago, 2013b, "Scaffolded memory and metacognitive feelings", in *Review of Philosophy and Psychology*, 4, 1: 135-152.
- Baars, Bernard J., 1997, *Inside the theatre of consciousness, the workspace of the mind*, Oxford University Press, Oxford.
- Bach, Dominik R. and Raymond J. Dolan, 2012, "Knowing how much you don't know: a neural organization of uncertainty estimates", in *Nature Reviews Neuroscience*, 13, 8: 572-586.
- Bjork, Rorbert A., 1999, "Assessing our own competence: Heuristics and illusions", in Daniel Gopher and Asher Koriati (eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application*, 435-459. MIT Press, Cambridge, MA.
- Brown, Alan S., 2003, "A review of the déjà vu experience", in *Psychological bulletin*, 129, 3: 394-413.
- Brown, Roger and David McNeill, 1966, "The 'tip of the tongue' phenomenon", in *Journal of verbal learning and verbal behavior*, 5, 4: 325-337.
- Byrne, Alex, 2005, "Introspection", in *Philosophical Topics*, 33, 1: 79-104.
- Carruthers, Glenn, 2012, "A metacognitive model of the sense of agency over thoughts", in *Cognitive Neuropsychiatry*, 17, 4: 291-314.
- De Neys, Wim, 2012, "Bias and conflict", in *Perspectives on Psychological Science*, 7, 1: 28-38.
- De Neys, Wim, Sofie Cromheeke, and Magda Osman, 2011, "Biased but in doubt: Conflict and decision confidence", in *PLoS ONE*, 6, 1: e15 954+.
- de Sousa, Ronald, 2008, "Epistemic feelings", in Georg Brun, Ulvi Doğuoğlu, and Dominique Kuenzle (eds.), *Epistemology and emotions*, Ashgate.
- de Sousa, Ronald, 2011, *Emotional truth*, Oxford University Press, Oxford.
- Dehaene, Stanislas and Lionel Naccache, 2001, "Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework", in *Cognition*, 79, 1-2: 1-37.
- Dennett, Daniel C., 2005, *Sweet dreams: Philosophical obstacles to a science of consciousness*, MIT Press, Cambridge, MA.

- Dokic, Jérôme, 2012, "Seeds of self-knowledge: noetic feelings and metacognition", in Michael J. Beran, Johannes L. Brandl, Josef Perner, and Joëlle Proust (eds.), *Foundations of metacognition*, Oxford University Press, Oxford: 302-321.
- Dokic, Jérôme and Paul Égré, 2009, "Margin for error and the transparency of knowledge", in *Synthese*, 166, 1: 1-20.
- Dokic, Jérôme and Jean-Rémy Martin, 2012, "Disjunctivism, hallucinations, and metacognition", in *Wiley Interdisciplinary Reviews: Cognitive Science*, 3, 5: 533-543.
- Efklides, Anastasia, 2002, "Metacognitive experiences in problem solving", in Anastasia Efklides, Julius Kuhl, and Richard M. Sorrentino (eds.), *Trends and prospects in motivation research*, Springer Netherlands, 297-323.
- Efklides, Anastasia and Alexandra Touroutoglou, 2010, "Prospective memory failure and the metacognitive experience of 'blank in the mind'", in Anastasia Efklides and Plousia Misailidi (eds.), *Trends and prospects in metacognition research*, Springer US: 105-126.
- Evans, Gareth, 1982, *The varieties of reference*, Oxford University Press, Oxford.
- Evans, Jonathan St. B. T. and Keith E. Stanovich, 2013, "Dual-process theories of higher cognition: Advancing the debate", in *Perspectives on Psychological Science*, 8, 3: 223-241.
- Finn, Bridgid, 2008, "Framing effects on metacognitive monitoring and control", in *Memory & Cognition*, 36, 4: 813-821.
- Flavell, John H., 1979, "Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry", in *American Psychologist*, 34, 10: 906-911.
- Fleming, Stephen M. and Raymond J. Dolan, 2012, "The neural basis of metacognitive ability", in *Philosophical transactions of the Royal Society of London Series B, Biological sciences*, 367, 1594: 1338-1349.
- Frankish, Keith, 2010, "Dual-process and dual-system theories of reasoning", in *Philosophy Compass*, 5, 10: 914-926.
- Goldinger, Stephen D. and Whitney A. Hansen, 2005, "Remembering by the seat of your pants", in *Psychological Science*, 16, 7: 525-529.
- Goldman, Alvin I., 2002, *Pathways to knowledge: Public and private*, Oxford University Press, Oxford.
- Gopnik, Alison, 1998, "Explanation as orgasm", in *Minds and Machines*, 8, 1: 101-118.
- Gopnik, Alison, 2000, "Explanation as orgasm and the drive for causal knowledge: The function, evolution, and phenomenology of the theory formation system", in Frank C. Keil and Robert A. Wilson (eds.), *Explanation and cognition*, MIT Press, Cambridge, MA: 299-323.
- Halamish, Vered, Shannon McGillivray, and Alan D. Castel 2011, "Monitoring one's own forgetting in younger and older adults", in *Psychology and aging*, 26, 3: 631-635.
- Hart, J.T., 1965, "Memory and the feeling-of-knowing experience", in *Journal of educational psychology*, 56, 4: 208-216.
- Hookway, Christopher, 2003, "Affective states and epistemic immediacy", *Metaphilosophy*, 34, 1-2: 78-96.

- Hookway, Christopher, 2008, "Epistemic immediacy, doubt and anxiety: On a role for affective states in epistemic evaluation", in Georg Brun, Ulvi Doguoğlu, and Dominique Kuenzle (eds.), *Epistemology and emotions*, Ashgate: 51-65.
- İnan, İlhan, 2012, *The philosophy of curiosity*, Routledge, London.
- James, William, 1890, *The principles of psychology*, Macmillan & Co, London.
- Johnson, Marcia K., Shahin Hashtroudi, and D. Stephen Lindsay, 1993, "Source monitoring", in *Psychol Bull*, 114, 1: 3-28.
- Koriat, Asher, 1993, "How do we know that we know? The accessibility model of the feeling of knowing", in *Psychological review*, 100, 4: 609-639.
- Koriat, Asher, 2000, "The feeling of knowing: Some metatheoretical implications for consciousness and control", in *Consciousness and Cognition*, 9, 2: 149-171.
- Koriat, Asher, 2006, "Metacognition and consciousness", in Philip David Zelazo, Moris Moscovitch, and Evan Thompson (eds.), *Cambridge handbook of consciousness*, Cambridge University Press, Cambridge.
- Koriat, Asher., Rorbert A. Bjork, Limor Sheffer, and Sarah K. Bar, 2004, "Predicting one's own forgetting: the role of experience-based and theory-based processes", in *Journal of experimental psychology General*, 133, 4: 643-656.
- Koriat, Asher. and Ravit Levy-Sadot, 2001, "The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing", in *Journal of experimental psychology Learning, memory, and cognition*, 27, 1: 34-53.
- Koriat, Asher. and Ravit Nussinson, 2009, "Attributing study effort to data-driven and goal-driven effects: implications for metacognitive judgments", in *Journal of experimental psychology Learning, memory, and cognition*, 35, 5: 1338-1343.
- Krug, Kevin, 2007, "The relationship between confidence and accuracy: Current thoughts of the literature and a new area of research", in *Applied Psychology in Criminal Justice*, 3, 1: 7-41.
- Kusumi, Takashi, 2006, "Human metacognition and the déjà vu phenomenon", in Kazuo Fujita and Shoji Itakura (eds.), *Diversity in cognition: Evolution, development, domestication, and pathology*, Kyoto University Press, Kyoto: 302-314.
- Levin, Daniel T. (ed.), 2004, *Thinking and seeing: Visual metacognition in adults and children*, MIT Press, Cambridge, MA.
- Litman, Jordan, Tiffany Hutchins, and Ryan Russon, 2005, "Epistemic curiosity, feeling-of-knowing, and exploratory behaviour", in *Cognition & Emotion*, 19, 4: 559-582.
- Loewenstein, George, 1994, "The psychology of curiosity: A review and reinterpretation", in *Psychological bulletin*, 116, 1: 75.
- Lorini, Emiliano and Cristiano Castelfranchi, 2007, "The cognitive structure of surprise: looking for basic principles", in *Topoi*, 26, 1: 133-149.
- Loussouarn, Anna, Damien Gabriel, and Joëlle Proust, 2011, "Exploring the informational sources of metaperception: The case of change blindness blindness", in *Consciousness and Cognition*.

- Luna, Karlos and Beatriz Martín-Luengo, 2012, "Confidence-accuracy calibration with general knowledge and eyewitness memory cued recall questions", in *Applied Cognitive Psychology*, 26, 2: 289-295.
- Mangan, Bruce, 1993, "Taking phenomenology seriously: The 'fringe' and its implications for cognitive research", in *Consciousness and Cognition*, 2, 2: 89-108.
- Mangan, Bruce, 2000, "What feeling is the 'feeling of knowing?'" in *Consciousness and cognition*, 9, 4: 538-544.
- Mangan, Bruce, 2001, "Sensation's ghost: The non-sensory 'fringe' of consciousness", in *Psyche*, 7, 18.
- Matthen, Mohan, 2005, *Seeing, doing, and knowing*, Oxford University Press, Oxford.
- Metcalf, Janet, 1993, "Novelty monitoring, metacognition, and control in a composite holographic associative recall model: implications for korsakoff amnesia", in *Psychological review*, 100, 1: 3-22.
- Michaelian, Kourken, 2011, "Generative memory", in *Philosophical Psychology*, 24, 3: 323-342.
- Michaelian, Kourken, 2012a, "Metacognition and endorsement", in *Mind & Language*, 27, 3: 284-307.
- Michaelian, Kourken, 2012b, "(Social) metacognition and (self-)trust", in *Review of Philosophy and Psychology*, 3, 4: 481-514.
- Michaelian, Kourken, 2013, "The information effect: Constructive memory, testimony, and epistemic luck", in *Synthese*, 190, 12: 2429-2456.
- Mitchell, Karen. J. and Marcia. K. Johnson, 2009, "Source monitoring 15 years later: what have we learned from fMRI about the neural mechanisms of source memory?", in *Psychological bulletin*, 135, 4: 638-677.
- Moore, George Edward, 1962, *Propositions: Some main problems of philosophy*, Collier.
- Moran, Richard, 2001, *Authority and estrangement*, Princeton University Press, Princeton.
- Morton, Adam, 2010, "Epistemic emotions", in Peter Goldie (ed.), *The Oxford handbook of philosophy of emotion*, Oxford University Press, Oxford: 385-399.
- Morton, Adam, 2013, *Emotion and imagination*, Polity, Cambridge.
- Moulin, C. and C. Souchay, 20xx, "Epistemic feelings and memory", in *Sage handbook of applied memory*.
- Nagel, Jennifer, 2010, "Epistemic anxiety and adaptive invariantism", in *Philosophical Perspectives*, 24, 1: 407-435.
- Nelson, Thomas O., 1999, "Cognition versus metacognition", in Robert Sternberg (ed.), *The nature of cognition*, MIT Press, Cambridge, MA: 625-641.
- Nelson, Thomas O. and Louis Narens, 1990, "Metamemory: a theoretical framework and new findings", in *Psychology of Learning and Motivation*, 26: 125-173.
- Norman, Elisabeth, Mark C. Price, and Simon C. Duff, 2010, "Fringe consciousness: A useful framework for clarifying the nature of experience-based metacognitive feelings", in Anastasia Efklides and Plousia Misailidi (eds.), *Trends and prospects in meta-cognition research*, Springer, US: 63-80.

- Oppenheimer, Daniel M., 2008, "The secret life of fluency", in *Trends in Cognitive Sciences*, 12, 6: 237-241.
- Origi, Gloria, 2008, *Qu'est-ce que la confiance?*, Vrin, Paris.
- Perfect, Timothy J., 2004, "The role of self-rated ability in the accuracy of confidence judgements in eyewitness memory and general knowledge", in *Applied Cognitive Psychology*, 18, 2: 157-168.
- Proust, Joëlle, 2001, "A plea for mental acts", in *Synthese*, 129, 1: 105-128.
- Proust, Joëlle, 2006, "Agency in schizophrenics from a control theory viewpoint", in Wolfgang Prinz and Natalie Sebanz (eds.), *Disorders of volition*, MIT Press, Cambridge: 87-118.
- Proust, Joëlle, 2007, "Metacognition and metarepresentation: is a self-directed theory of mind a precondition for metacognition?", in *Synthese*, 159, 2: 271-295.
- Proust, Joëlle, 2008, "Epistemic agency and metacognition: An externalist view", in *Proceedings of the Aristotelian Society*, 108, 1: 241-268.
- Proust, Joëlle, 2009a, "Is there a sense of agency for thought?", in L. O'Brien and M. Soteriou (eds.), *Mental actions and agency*, Oxford University Press, Oxford: 253-279.
- Proust, Joëlle, 2009b, "The representational basis of brute metacognition: A proposal", in Robert W. Lurz (ed.), *The philosophy of animal minds*, Cambridge University Press, Cambridge: 165-183.
- Proust, Joëlle, 2012, "Metacognition and mindreading: One or two functions?", in Michael J. Beran, Johannes L. Brandl, Josef Perner, and Joëlle Proust (eds.), *Foundations of metacognition*, Oxford University Press, Oxford, 234-251.
- Proust, Joëlle, 20xx, *The philosophy of metacognition: Mental agency and self-awareness*, Oxford University Press, Oxford.
- Reder, Lynne M., 1988, "Strategic control of retrieval strategies", in *Psychology of Learning and Motivation*, 22: 227-259.
- Reder, Lynne M., 1996, *Implicit memory and metacognition*, Lawrence Erlbaum Associates.
- Reder, Lynne M., 1987, "Strategy selection in question answering", in *Cognitive Psychology*, 19, 1: 90-138.
- Reder, Lynne M. and Frank E. Ritter, 1992, "What determines initial feeling of knowing? Familiarity with question terms, not with the answer", in *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 3: 435-451.
- Rosenthal, David, 2000, "Consciousness and metacognition", in Dan Sperber (ed.), *Metarepresentation: Proceedings of the tenth Vancouver cognitive science conference*, Oxford University Press, Oxford: 265-295.
- Rosenthal, David, 2012, "Higher-order awareness, misrepresentation and function", in *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367, 1594: 1424-1438.
- Russell, Bertrand, 1921, *The analysis of mind*, George Allen & Unwin, London.

- Sainsbury, R.M., 1995, "Vagueness, ignorance, and margin for error", *The British Journal for the Philosophy of Science*, 46, 4: 589-601.
- Schacter, Daniel L. and Donna R. Addis, 2007, "The cognitive neuroscience of constructive memory: remembering the past and imagining the future", in *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362, 1481: 773-786.
- Schwartz, Bennett L. and Janet Metcalfe, 2011, "Tip-of-the-tongue (TOT) states: retrieval, behavior, and experience", in *Memory & Cognition*, 39, 5: 737-749.
- Silvia, Paul J., 2006, *Exploring the psychology of interest*, Oxford University Press, Oxford.
- Smith, J. David, 2009, "The study of animal metacognition", in *Trends in Cognitive Sciences*, 13, 9: 389-396.
- Smith, J. David, Wendy E. Shields, and David A. Washburn, 2003, "The comparative psychology of uncertainty monitoring and metacognition", in *Behavioral and Brain Sciences*, 26: 317-339.
- Spehn, Michaela K. and Lynne M. Reder, 2000, "The unconscious feeling of knowing: A commentary on Koriat's paper", in *Consciousness and Cognition*, 9, 2: 187-192.
- Strawson, Galen, 2011, "Cognitive phenomenology: Real life", in Tim Bayne and Michelle Montague (eds.), *Cognitive phenomenology*, Oxford University Press, Oxford: 285-325.
- Thompson, Valerie A., 2009, "Dual-process theories: A metacognitive perspective", in Jonathan St. B. T. Evans and Keith Frankish (eds.), *In two minds: Dual processes and beyond*, Oxford University Press, Oxford: 171-195.
- Thompson, Valerie A., Jamie A. Prowse Turner, and Gordon Pennycook, 2011, "Intuition, reason, and metacognition", in *Cognitive Psychology*, 63, 3: 107-140.
- Trout, J.D., 2002, "Scientific explanation and the sense of understanding", in *Philosophy of Science*, 69, 2: 212-233.
- Trout, J.D., 2007, "The psychology of scientific explanation", in *Philosophy Compass*, 2, 3: 564-591.
- Wessel, Jan R., 2012, "Error awareness and the error-related negativity: Evaluating the first decade of evidence", in *Frontiers in Human Neuroscience*, 6, 88.
- Whittlesea, Bruce W., 1993, "Illusions of familiarity", in *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 6: 1235-1253.
- Whittlesea, Bruce W. and Lisa D. Williams, 2000, "The source of feelings of familiarity: the discrepancy-attribution hypothesis", in *Journal of experimental psychology Learning, memory, and cognition*, 26, 3: 547-565.
- Whittlesea, Bruce W. and Lisa D. Williams, 2001, "The discrepancy-attribution hypothesis: I. the heuristic basis of feelings of familiarity", in *Journal of experimental psychology Learning, memory, and cognition*, 27, 1: 3-13.
- Whittlesea, Bruce W. and Lisa D. Williams, 1998, "Why do strangers feel familiar, but friends don't? A discrepancy-attribution account of feelings of familiarity", in *Acta Psychologica*, 98, 2-3: 141-165.
- Yeung, Nick and Christopher Summerfield, 2012, "Metacognition in human decision-making: confidence and error monitoring", *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367, 1594: 1310-1321.