
The Representational Structure of Feelings

Joëlle Proust

The word “feeling” denotes a reactive, subjective experience with a distinctive embodied phenomenal quality. Several types of feelings are usually distinguished, such as bodily, agentive, affective, and metacognitive feelings. The hypothesis developed in this article is that all feelings are represented in a specialized, non-conceptual “expressive” mode, whose function is evaluative and action-guiding. Feelings, it is claimed, are conceptually impenetrable. Against a two-factor theory of feelings, it is argued, in the cases of affective and metacognitive feelings, that background beliefs can circumvent feelings in gaining the control of action, but cannot fully suppress them or their motivational potential.

Keywords

Affective feelings | Affordance | Agentive feelings | Appraisal | Arousal | Bodily feelings | Comparator | Control | Cues | Evaluative | Expressive | Familiarity | Fluency | Formal object | Illusory feeling | Incidental feelings | Integral feelings | Intensity | Metacognitive or noetic feelings | Monitoring | Nonconceptual content | Predictive | Reactive | Resonance | Retrospective | Somatic marker | Transparency | Two-factor account | Valence

Author

Joëlle Proust

joelle.proust@ehess.fr

Ecole Normale Supérieure
Paris, France

Commentator

Iuliia Plushch

pliuschi@students.uni-mainz.de

Johannes Gutenberg-Universität
Mainz, Germany

Editors

Thomas Metzinger

metzinger@uni-mainz.de

Johannes Gutenberg-Universität
Mainz, Germany

Jennifer M. Windt

jennifer.windt@monash.edu

Monash University
Melbourne, Australia

1 Introduction

“Feeling” denotes a reactive, subjective experience with a distinctive embodied phenomenal quality and a formal object, which may or may not coincide with embodied experience. Feelings typically express affect and valence in sensation. “Reactive” means that feelings are closely associated with an appraisal of a present property or event. The term “reactive” is crucial. The term “feeling” is sometimes used to refer to a non-reactive, perceptual experience. For example, when one perceives an object through touch, it is common to say that “one feels one’s key in one’s pocket”. But “feeling”, in this context, does not refer to a reactive phenomenon.

It rather refers to the feedback of one’s own key-touching activity. This type of perceptual feeling is expected to result from one’s action and, hence, does not belong to the domain of reactive feelings. What is called the “formal object” (see [Kenny 1963](#)) of a feeling is the property in the triggering event that elicits the reactive feeling. For example, the formal object of fear is some threatening property detected in the perceptual field.

Feelings can be pleasant or aversive, strong or weak, short-lived or long-lasting, or have an arousing or depressing character. They motivate distinctive dispositions to act, whose

Glossary

Feeling	“Feeling” denotes a reactive, subjective experience with a distinctive embodied phenomenal quality and a formal object, which may or may not coincide with the embodied experience. Feelings typically express affect and valence in sensation.
Reactive	“Reactive” means that feelings are closely associated with an appraisal of a present property or event.
Formal object	“Formal object” of a feeling is the property in the triggering event that elicits the reactive feeling.
Metacognitive feelings	Metacognitive feeling are experienced while conducting a cognitive task: the agent may find the task easy or difficult, anticipate her ability or inability to conduct it. Once the task is completed, the agent may have the feeling of being right, or have a feeling of uncertainty about the outcome of her endeavour.
Affordance	Affordances are positive or negative opportunities, expressed in feelings: an affordance-sensing swiftly and non-reflectively motivates the agent to act in a particular way.
FS Affordance	FS Affordance _a [Place _a =here], [Time _a =Now/soon], [Valence _a =+], [Intensity _a =.8 (comparatively specified on a scale 0 to 1)], [motivation to act of degreed according to action programa].
Transparency	A mental state is transparent if, when it is activated, its intentional content is accessible to the subject who entertains it.
Incidental and integral feeling	Metacognitive feelings are called “incidental” when they are not based on valid cues about the cognitive task at hand, and hence, have no predictive value. They are called “integral” when they actually carry information about cognitive outcome.

urgency is entailed both by the feeling experience and the context in which it is experienced: feeling an intense pain disposes the person to promptly locate and remove the cause of the pain; except, for example, when it is self-inflicted, or when it is part of a ritual.

Most theorists of feelings agree that they are associated with—or, for those who identify emotions with conscious experiences¹ consist of—specialized, internally generated bodily sensations, such as an increase in heart rate, contractions or relaxations of the facial muscles, visceral impressions, tremors or tears, impulses to run away, etc. As will be seen below, some feelings, however, do not express emotions., i.e.,

¹ From the viewpoint of the somatic feeling theory of emotions, emotions can be explained as a somatic change caused by the perception, real or simulated, of a particular object. See [James \(1884, p. 190\)](#), and [Damasio \(1994, 2003\)](#). Other theorists of emotion, however, consider that the conscious experience of having an emotion includes propositional attitudes, and not only feelings. See sections 4 and 5 below. Moods are long-term affective states, and will not concern us here.

they are not affective. A feeling tends to be more explicitly felt as bodily when it has a body-related function; that is, the phenomenology makes the need to be served salient (feeling tired, feeling a pain in the joints) in order to motivate action. In affective feelings, in contrast, the bodily phenomenology tends to recede to the fringe of consciousness (feeling in love with *A*, feeling angry with *B*).² From this observation, it is easy to infer that types of feelings differ in their respective meanings: they in some sense *express what they are about*. In affective feelings, an experience of “feeling toward” is supposedly present: the emotion is felt as being about an object, a person, or a situation—the objects, rather than bodily sensations, are the focus of one’s emotional attention. Affective feelings also include mixed cases where one seems to both experience a strong bodily feeling at the same time as the intentional content that

² On this concept, see [Mangan \(1993, 2000\)](#) and [Reber et al. \(2002\)](#).

this feeling seems to refer to, as when Marcel Proust's narrator reports experiencing an acute pain in the chest when thinking about his beloved deceased friend, Madame de Guermantes.³ It is unclear whether metacognitive (also called noetic, or epistemic) feelings are affective or non-affective (see section 7 below). They are experienced while conducting a cognitive task: the agent may find the task easy or difficult, and may anticipate her ability or inability to conduct it. Once the task is completed, the agent may have the feeling of being right, or may have a feeling of uncertainty about the outcome of her endeavour. Take the case of a person who feels unable, presently, to remember what she had for dinner last night. Her feeling of not remembering is correlated with activity in a facial muscle, the corrugator supercilii (Stepper & Strack 1993). Her feeling, however, is not about her disposition to contract or relax this or that muscle, of which she is certainly unaware. It is, rather, about her present disposition to remember what she had for dinner. Epistemic feelings seem to be “feeling-toward” experiences, and have cognitive dispositions or contents as their object.

Descriptive phenomenology, however, does not offer in itself an account of the intentional structure of feelings. We need to understand how feelings in general gain their real or supposed aboutness, and how they relate to action-guidance as a function of context; i.e., we need to provide a functional analysis of feelings. Section 2 will begin to provide such an analysis, and will address a preliminary issue—namely, Do the phenomena that are usually called “feelings” share a property that makes them a natural kind? In section 3, the specific informational structure of feelings will be seen to account for their generic characteristics. Section 4 will clarify the account by way of addressing various objections. Section 5 will attempt to show that the proposed account fares better with experimental evidence than a cognitivist account of affective and metacognitive feelings. Section 6 will examine whether or not metacognitive feelings have an affective valence.

2 Are feelings a natural kind?

Paul Griffiths has claimed that *emotions* do not constitute a natural kind, in the sense that they do not form “a category about which we can make inductive scientific discoveries” (2004, pp. 901–911). One can agree with latter claim, however, without concluding that *feelings* do not constitute a natural kind. First, feelings are not only affective ingredients in emotional awareness. Some feelings, such as feeling cold or sick, or feeling that one is acting, have nothing to do with affective episodes. Second, there are evolutionary reasons to distinguish, within emotions, two classes of subjective appraisals. Emotion theorists usually contrast *feelings* expressed in primary emotions—fear, anger, happiness, sadness, surprise, and disgust—with *various appraisals cum conative dispositions* associated with higher cognitive emotions, such as envy, guilt, pride, shame, loyalty, vengefulness, and regret. The first are phylogenetically and ontogenetically prior to cognitions. They belong to the ancient limbic system, which is present in some form in most animals. A quick route from the retinal image to the amygdala through the thalamus allows affective information to control behavior (see LeDoux 1996). Primary feelings are thus triggered independently of concept possession and motivate specific responses. Secondary affective experiences, in contrast, might have evolved on the basis of social constraints in relation to cooperative action among humans. Indeed (with the possible exception of pride and shame) they are not present in nonhuman *primates*.⁴ They activate newer brain structures; they require concept possession, depend on background beliefs, and do not generate characteristic behaviors. Finally, primary feelings are clearly embodied, while secondary emotions seem to have no proprietary somatic markers. An interesting idea, suggested by Jesse Prinz (2004, p. 95), is that the facial or somatic correlate of secondary emotions, *when they have one*, involves a blend of the somatic markers for primary feelings.

³ See the analysis of this example in Goldie (2002), p. 56.

⁴ On this contrast, see Frank (1988), Griffiths (1997), and Prinz (2004, pp. 82–83). On whether they qualify as emotions, see Ekman (1992).

In summary: emotions differ, among other things, because of the unequal role that feelings have in the two classes of emotions just discussed. The wider scope of feelings, when understood as “reactive, subjective experiences with a distinctive embodied phenomenal quality”, seems to be more unified than emotions, and making feelings seem like plausible candidates for a natural kind.

We need, however, to turn this tentative definition into a general functional characterization that presumably holds for all feelings (beyond affective ones) and only for them. Here is a proposal: feelings constitute the sensitive part of predictive and retrospective processes of non-conceptual evaluation of one’s own and others’ well-being and actions. Being essentially evaluative, feelings are always the output of a comparator: in other terms, they are crucial *monitoring* ingredients in self-regulated adaptive control systems. In such systems, the specific function of a feeling consists in detecting how much a current *observed* value of a parameter *deviates from its expected value*, on one or several dimensions relevant to survival (see Carver & Scheier 2001). Their formal object, when they have one,⁵ (such as being afraid of the bear in front of me) cannot be analyzed independently of the monitoring function they serve within a specialized control loop.⁶ Relevance to well-being, however, extends to bodily condition, goal achievement, and availability of preferred goods of all kinds (food, partner, social status). The relevant dimensions of variation that feelings track may accordingly be of a sensory, proprioceptive kind (feeling thirsty, cold, etc.), social-affective (feeling angry), or agentive (goal-related). Goal achievement, however, involves either epistemic or instrumental success, respectively generating epistemic feelings (feeling interested, bored, epistemically uncertain) and agentive feelings (feeling of happiness, of agentive confidence, of ownership of one’s action, etc.). Feelings, in summary, are the outcomes of comparators in a con-

trol loop; they carry non-conceptual information about how much one’s present condition deviates from the expected condition. From a functional viewpoint, they form a natural kind insofar as their function is to indicate a comparative outcome through a dedicated embodied experience.

Note, however, that there are comparators that trigger no feelings at all: these non-sensitive comparators may either work outside consciousness (for example, error signals driving immediate correction⁷, not to mention comparators that work at the cell level), or they can take concepts as their input, rather than reacting to percepts or situations (for example comparators of currency or of educational value).

As far as feelings are concerned, they are directly related to a presently-perceived context (or an imagined or remembered context, but in a “present-like”, indexical mode): one can feel too hot, too cold, or too tired (or feel “OK”, which usually means a tolerable deviation from the expected value). One can feel the fright one has had, even after the frightening event has ended. The outcome of a feeling-based appraisal, from a functional viewpoint, has to consist in some disposition to act that is adaptive, relative to the input to which the feeling is a reaction. Granting that feelings, as sensitive comparators in a control system, form a natural kind, there should be common properties cutting across the various types listed above. In fact we find three types of functional relations between feelings of a given kind and the associated disposition to act. First, feelings, according to their embodied valence, typically determine actions of approach or of avoidance. Some dictate caution, others boldness. Some encourage self-restraint, others self-assertion. Fear promotes a flight tendency, hunger a tendency to approach food. Second, they have a specific orientation in time: some feelings have a *predictive* function, and thus induce a behavior that is based on contingencies to be *further* displayed in the present context. For example, fear, when directed at a possible danger, increases the readiness to flee in case the danger concretizes.

⁵ As observed by Goldie (2009), some feelings, for example, [feeling anxious] or [feeling depressed], seem to lack a formal object, which is typically the case with moods. As indicated above, moods will not be discussed in this article.

⁶ Bechara et al. (2000) make it clear that the somatic marker theory applies to action, whether it engages affects or not.

⁷ see Logan & Crump 2010 and Nieuwenhuis et al. 2001

Others have a *retrospective* function, and induce corrections to the commands one has previously used, or to one's previous preferences. For example, feeling nauseous *after* food ingestion induces food avoidance, i.e., a change in the agent's preferences. In contrast, feeling disgust at the sight of some food may prevent the agent from approaching it. A subset of feelings, such as feeling happy, have both temporal orientations. Third, according to their embodied dynamics and intensity (which is called their "level of arousal"), feelings can provoke an elevation in the energy available to the system: they provoke excitement, agitation, power in the coming response; or, on the contrary, they may have a soothing effect and diminish the tendency to act.

One major functional property of feelings, from the viewpoint of information extraction and use, is that they can very rapidly extract and synthesize multiple cues from perception. This rapidity is a consequence of the automatic and encapsulated character of the control mechanism whose output they express. Feelings are automatically triggered by a specific type of input (which is the definition of informational encapsulation).⁸ Automaticity is associated with feelings being inescapable, at least for those feelings that have been allowed to develop within a culture, granting normal development.⁹ The mechanism that generates somatic, noetic, or affective feelings from inputs (perceptual, imaginative, or memorial) does not require one to have specific beliefs or intentions.¹⁰ Informational encapsulation explains why transitive feelings persist when the agent finds out that the situation is different from what she thought to be the case. Just as an optical illusion such as the Müller-Lyer effect does not immediately dissipate when it turns out that the segments are equal, a feeling of anger does not disappear as soon as the agent realizes that its formal object is not exemplified.

⁸ Automaticity in appraisal is central to Ekman's analysis of primary emotions (1992). See also Griffiths (1997), Prinz (2004), and Zajonc (1980). On informational encapsulation, see Fodor (1983).

⁹ For example, fearlessness in the presence of danger may result from a disturbed childhood.

¹⁰ Some affective feelings, however, can be intentionally controlled in the long run, through cultural learning. See Murata et al. (2013).

Automaticity and informational encapsulation seem also to characterize agentive feelings (see Pacherie 2008). Feelings generated in the course of a physical action come in two varieties: generalized or specialized. Some, such as feelings of agency, of initiation of action, of ownership and of motor control, are indicators monitoring action in progress: they concern "who" is performing the action, and "how" the action is being conducted (see Proust 2000). Others concern the evaluation of an action in one's own repertoire: a professional carpenter or an experienced musician, for example, have feelings telling them if an action sequence (whether their own or another agent's) in this repertoire sounds or looks right, even before they identify why they have this feeling. These feelings are also the outputs of a comparison between motor anticipations and observed properties of the action (a "forward model of action" supposedly stores the expected values of crucial parameters; Wolpert et al. 2001). They can predict the likelihood of completing an action (when the question arises, in difficult or non-routine cases), or evaluate—on-line or in retrospect—how swiftly, effortlessly, or unhesitatingly an action was performed. Agentive feelings thus have an essential role in regulating the fundamental properties of physical actions, such as the quality of the outcome,¹¹ and the ownership of the action.¹²

Noetic feelings, finally, are functionally similar to somatic, affective, and agentive feelings—although their evolutionary pattern seems to be different from the other three kinds. While most organisms have proprioceptive, affective, and motor control, and hence, presumably, somatic, affective, and agentive feelings, few are able to control their cognitive decisions through metacognitive feelings (see Beran et al. 2012 and Proust 2013). The latter are generated when trying to perceive, to remember, or to plan a cognitive task (in particular, when trying to plan how long to study material in order to

¹¹ Non-conscious error signals can also guide corrective steps, without the agent noticing them.

¹² Pat Haggard et al. (2002) have demonstrated the crucial role of the temporal binding between felt initiation of action and output in the sense of being the agent of an action. See, among other articles, Haggard et al. (2002).

master it).¹³ They are also relied upon when trying to reason or to solve a problem; when conversing, feelings of effort, and of informativeness, are monitored by speakers and hearers in order to maintain a common level of relevance. Like other feelings, they have two distinctive temporal orientations. Some have a predictive function. A feeling of knowing (FOK) may arise when trying to remember an item—for example a proper name—that one has not yet retrieved: having a strong FOK reliably predicts that one will finally retrieve the searched content (Koriat & Levy-Sadot 2001). A feeling of having a name on the tip of one's tongue (TOT) both signals the fact that a word is not presently available, and, according to its onset, valence, and intensity, whether it is worth or not worth pursuing one's effort to retrieve it (see Brown 1991 and Schwartz et al. 2000). Feelings of fluency are the sense of ease of processing one may feel or fail to feel when attempting to perceptually discriminate objects with a given property, or to retrieve items from episodic or semantic memory. A feeling of familiarity is particularly salient, in human adults, when no further fact about the target can be retrieved. It offers useful information about the epistemic status of the target: that it is not new, but nevertheless not fully recognized. A feeling of familiarity, then, motivates, among others, an attempt to recognize what or who a target is. Other metacognitive feelings have a retrospective function. When a name is retrieved, a feeling of rightness (FOR) motivates the agent to consider her response the expected one.¹⁴ Various feelings of uncertainty, based on fluency, coherence, plausibility, informativeness, or relevance, also have retrospective functions: their valence and intensity tell the agent whether she should accept or reject a cognitive outcome. These parameters are expressed through specialized somatic markers, such as increased activity in the facial muscle involved in smiling, the zygomaticus major—for positive valence—or the corrugator supercilli (involved in frowning)—for negative valence (Winkielman & Cacioppo 2001).

Taken together, these considerations are compatible with the view that somatic, agentive, metacognitive, and “primary” affective feelings, even if they differ in their formal objects, form a natural kind. Our attempt above at a functional characterization focused on the general relations of feelings to inputs, outputs, and mediating evaluative mechanisms. From this characterization, it emerges that feelings are gradients in comparators that are felt subjectively, rather than being propositional states describable in analytic, objective terms. These observations, however, suggest that, in order to express a specialized and fine-tuned reactivity to one or several formal objects, and to motivate adapted behaviors, in order to be remembered and conveyed to others feelings must have their own representational format. We now turn to the following question: What is the structure of the information that is extracted and expressed in a feeling?

3 What kind of information do feelings express?

The above question is important for clarifying the relation of feelings both to their formal object, when they have one, and to the action that they motivate. In the case of metacognitive feelings (M-feelings), the difficulty is particularly pregnant: it stems from the fact that, if we grant that M-feelings do not require concept possession to be felt, then it is unclear how their formal object should be construed: What are they about? Let us take a feeling of uncertainty, felt while trying to remember a proper name. Is this feeling about a memory *state*, or about a *disposition* to retrieve a proper name? If a feeling is about a memorial state or a disposition, its intentional content needs to include concepts of memory, of correctness, and of uncertainty. Empirical evidence, however, demonstrates that animals with no mindreading ability, and hence that are deprived of concepts of perception or of memory, are able to monitor their perception and memory as reliably as humans do.¹⁵ Furthermore, human children, from

¹³ This prediction involves judgments of learning (JOL). See Koriat & Ackerman (2010).

¹⁴ On FORs, see Thompson et al. (2011).

¹⁵ Rhesus monkeys have been found to opt out of more or less challenging perceptual or memory trials as a result of trial difficulty. For a

early on, are sensitive to the contrast between familiar and unfamiliar faces and environments. This supports our claim above: one can feel cold or anxious or uncertain without having the corresponding concepts of those feelings. A propositional format does not seem to apply to feelings in general.¹⁶

How do feelings fulfill their particular embodied, subjective way of representing—a mode we will call the “expressive mode”? The broadly functional characterization given above provides useful clues. Expressive representations comprise exclusively non-conceptual, perceptual, and evaluative (gradient- and valence-based) elements, which taken together express a subjective relation to the environment (internal or external) and a given tendency to act. It should be emphasized, however, that adult humans can obviously entertain *simultaneously* expressive and conceptual representations. The present hypothesis, in conformity with the literature on dual-processing, is that the expressive system processes information and influences decisions on the basis of its own narrow range of associations and norms; while the conceptual system takes advantage of background beliefs and inferential reasoning to make decisions in light of a broader set of norms. Let us take the case of an agent feeling joy after having won the lottery. A human adult normally has [lottery] in her conceptual repertoire, along with some of the inferences that can be made on its basis. However, the agent’s reactivity to the winning event falls under the expressive mode of representation, because this is the mode in which evaluation of the opportunities is conducted. This feeling representation presumably enlightens and orients the concept-based reasoning that can be conducted concerning the same event, such as wondering how to spend the money, or whether quitting her job is a good idea. We propose to call “affordance-sensing” the information that a feeling expresses. Affordances are positive or negative opportunities, expressed in feelings: an affordance-sensing swiftly and non-reflectively

motivates the agent to act in a particular way. Departing somewhat from Gibson’s use of this term within his ecological theory of perception, “affordance” is used here to refer to a non-conceptual and entirely subjective appraisal of the environment by the agent: an affordance is a perceived utility, which can be positive (something to approach and grasp) or negative (something to avoid and from which to flee).¹⁷

The corresponding representation has an indexical structure, because it has an essential relation to an occurrent represented property. Indexicality, however, has to be understood here in a non-referential sense. What is indexed is an *occurrent* (relational) affordance, rather than an individual event or object. Here is our proposal for what a given feeling structure (FS) looks like:

- FS Affordance_a [Place_a=here], [Time_a=now/soon], [Valence_a=+], [Intensity_a=s (comparatively specified on a scale 0 to 1)], [motivation to act of degree_a according to action program_a].

The subscript “_a” is meant to indicate that all the elements that have this subscript are representational cues, i.e., ingredients, in present affordance-sensing *a*. Note that the strength (or degree) of the motivation to act does not depend only on the fitness significance, i.e., on the valence and intensity of the affordance. Other factors, such as the physical condition of the agent and her prior arousal level (her mood) also modulate her motivational level (Schwarz & Clore 2007). The specification of the location of the affordance may vary, depending on the way the feeling was generated, but indexicality and reactivity suggest that the relevant affordance is often sensed to occur where the feeling is experienced. As will be seen later, however, M-feelings do not involve a specification of place.

The feeling structure proposed above includes somatic markers, even if they are not

summary of the results and a methodological discussion of their significance, see Beran et al. (2012), Chapter 1, and Proust (2013), Chapter 5.

¹⁶ For a defense of emotional representations as nonconceptual and action oriented, see Griffiths & Scarantino (2009).

¹⁷ See Proust (2009, 2013). Prinz (2004) briefly discusses this idea in connection with the intentional content of emotions (p. 228). See also Griffiths & Scarantino (2009): in emotion, “the environment is represented in terms of what it affords to the emoter in the way of skillful engagement with it.”

made explicit: these markers are the substrates for the information of valence and intensity. This information is carried by neural activations and associated bodily changes, such as a sudden sensation of pleasant muscle relaxation, or of unpleasant muscle contraction, or of visceral contractions associated with fear. Intensity of affordance, i.e., the arousal produced by a feeling, is also felt through the comparative amount of bodily reactivity to the affordance. These somatic markers, as emphasized above, are themselves part of a monitoring system designed to predict and assess one's relations to the environment along the relevant dimensions listed above (agency, individual and social well-being, preferences, and metacognition).

Let us consider further how to read the feeling structure given above. It is meant to reflect not only what is presently felt, but also what is stored in memory when a feeling is experienced, what can be imagined, and what can be conveyed to others in expressive behavior. The central idea is that feelings sensitively express *a subjective, embodied relation to an opportunity* in an input from the environment (understood in a broad sense as including external and bodily properties relevant to well-being). This primitive intentional relation is best captured by the term *affordance-sensing*. Feelings express this affordance as their focus (or formal object), along with its graded valence—ranging from very unpleasant to very pleasant—and with its intensity gradient, which ranges from small to large.¹⁸

As often emphasized, reactivity to an affordance occurs very rapidly in a processing sequence—even before the perceptual processing has been completed—and well before a concept-based judgment can be made (see Dolan 2002, p. 1191; Griffiths 1997, pp. 77; LeDoux 1996, pp. 174; Prinz 2004, pp. 60, and Zajonc 1980, pp. 153). This suggests that an alternative,

evaluative informational system screens the input with its own independent memorial structures.¹⁹

An affordance does not need to have an objective counterpart to be sensed, i.e., for a feeling to arise: it is enough that the agent anticipates it (even wrongly), imagines it, or remembers it, for the corresponding feeling to be expressed. A feeling, thus, does not presuppose a conceptual appraisal of the context, but rather it indexes in an embodied way a direct evaluative registration. Given that an affordance does not aim at characterizing the world, one cannot say, when the expressed affordance has no objective counterpart, that a feeling “misrepresents” the world as having a given affordance, or reciprocally that an existing affordance was “missed” by the agent when the latter failed to detect it. For misrepresentation to occur, a system must be equipped to attribute properties to individual objects, that is, it must be able to apply concepts. The expressive system, however, does not refer to objects as independent entities. Hence, affordance is not literally what a feeling is about, because aboutness presupposes that what is represented is independent from the representational system. Being relational, affordances cannot be grasped independently of the experience of a sensitive agent. When saying that a feeling “expresses” an affordance, we mean that it “resonates” to it (or that it monitors it). Resonance is a neural-somatic reactivity: it carries indexical and evaluative information, but it does not refer to the world or attempt to describe it.

It is possible, however, to objectively characterize what a feeling functionally refers to, and to pinpoint cases of misrepresentation, by re-describing the feeling structure above in non-subjective, non-evaluative propositional terms. Taking advantage of her perceptual and background beliefs, the agent can claim to have mistaken a piece of wood for a snake, for example,

¹⁸ For a review of the theories of valence, see Prinz (2004), Ch. 7. Prinz takes valence to be a different determinate experience in each feeling. On valence as determined by overall value, from a consumer semantics viewpoint, rather than as an experience of pleasure/displeasure, see Carruthers (2011), pp. 127–130. This view, however, does not build on the nonconceptual information being felt, but rather on its being represented “in an abstract and amodal way”, which, nevertheless, is motivating.

¹⁹ These expressive representations do not require a system to have the capacity to form propositional representations. They are close to what Strawson called “a featural representational system”, allowing an animal to navigate with no propositional thinking (1959). On the comparison between the two representational modes, see Proust (2013). The question of the penetrability of feelings by propositional thought is explored below, in section 5.

and to make explicit that there is no reason to be afraid of a piece of wood.

Our analysis of FS helps us to clarify why “feeling one’s keys in one’s pocket” does not belong to reactive feelings. Recognizing through touch the object in one’s pocket as being one’s keys, or merely having a proprioceptive experience in fact caused by one’s keys, are two ways of perceiving one’s keys, involving respectively cognitive and sensory proprioception. But neither needs as such to involve an affordance of a given intensity and valence. In contrast, let us suppose that the perceiver believes wrongly that she has forgotten her home keys, which are in some distant location, and will not be able to get back home. Feeling her keys in her pocket immediately triggers a positive affordance, opening up the field of possible actions.

4 Questions and objections

The present proposal raises a number of additional questions and objections. Let us start with the most radical objection.

4.1 Are feelings representations?

Granting that feelings, affective or not, can be pure “physical effects of objects on the nerves”, in [William James’](#) terms (1890, vol. 2, p. 458), they do not need to have any genuine representational value. James invites us to take the case either of purely somatic feelings or of objectless emotions when they are generated by a pathological condition—such as the precordial catch syndrome (PCS) which is a feeling of pain in the chest that usually goes away without treatment, but can lead the victim to think he or she is suffering a heart attack. In this case, the emotional experience of dread, [James](#) says, is “nothing but the feeling of a bodily state, and it has a purely bodily cause” (1890, vol. 2, p. 459). From this, one might conclude that a feeling is a merely peripheral phenomenon: it does not have a function to represent, nor does it express anything in particular. What can be said, in response, is, first, that feelings have a crucial evaluative function, which they perform thanks to their expressive structure. In PCS, the pa-

tient’s experience of dread has valence and intensity, expressed through sudden breathlessness, chest constriction, blurred vision, tingling sensations in the skin, an elevated heart beat, and a disposition to crouch. These feelings are not only a matter of sensory “peripheral” experience: they are also used by the patient to collect her existing Bayesian correlations, and to monitor with their help the present affordance expressed. A second illustration of the representational nature of feelings is that they can arise in the absence of the sensory basis they seem to have. For example, illusory feelings of being touched—a reactive somatosensory feeling about a change occurring on one’s body surface—can be created by manipulating the coherence of the intermodal inputs from vision, touch, and proprioception. In the so-called “rubber-hand illusion”, participants feel that their hand is being touched with a paint brush, when in fact it is an artificial hand, not theirs, that they see being touched. They also, after a while, “feel as if their (real) hand is turning ‘rubbery’” (see [Botvinick & Cohen 1998](#)). This experiment is evidence that feelings are informational states, which monitor inputs, and, in extreme cases like this, cause the brain to try to reconcile contradictory multimodal input. In the proposed interpretation, however, seeing one’s hand being touched is a reactive feeling, while actively touching an object generates a percept—which plays quite a different role in cognition.

4.2 What does “resonating to an affordance” mean?

Second, speaking of “subjective resonance” to an affordance (see the discussion of how a feeling “resonates” to an affordance in section 3 above) may look improperly metaphorical.²⁰ This is meant, however, to mark the difference between feeling and perceiving. While percepts allow recognition and identification of external objects and properties, feelings express specific affordances in a perceived, imagined, or remembered situation. For example, one can feel cold right now, or simulate being cold when

²⁰ In a similar vein, [William James](#) writes that, in emotions, “the whole organism is a sounding board” (1890, vol. 2, p. 450).

planning a polar trip; one can remember how angry, or bored one was in a given episode and context. Feelings give agents prompt access to the relevant features of a new situation through sensed changes in their experience. Importantly, resonance is also an apt term for empathy, i.e., for the propagation of feelings from an agent to an onlooker, based on expressive behavior (Decety & Meyer 2008; Dezechache et al. 2013). Brain imagery suggests that the perception of pain in another individual largely overlaps with the regions activated when experiencing pain oneself (Jackson et al. 2005). Such empathy, in the present proposal, exemplifies how a feeling structure can be communicated through a set of congruent behavioral cues associated with a given affordance (here a painful stimulus), with a valence and intensity that are bodily conveyed.

4.3 Non-conceptual content as a common feature of feelings and percepts

Third, one might object that a common feature of feelings and percepts is that they include non-conceptual contents. This is true; but notice the difference between the two types of non-conceptual content: while non-conceptual ingredients in perception are related to objective, external contrastive cues such as shapes, edges, colors, volumes, and auditory patterns, which can be static or dynamic, but are always purely descriptive, non-conceptual contents in feelings only include evaluative states, which combine the general type of the affordance, its valence, its intensity, the proper action program, where all constituents are “bodily marked”, i.e., expressed through somatic markers. Therefore we cannot say that feelings “perceive” affordances, for this would suppose either that feelings have direct sensory access to the world—which they don’t, for they extract their inputs from sensory perception—or that they have direct sensory access to the body, which they don’t have either—feelings are the subjective counterpart of bodily changes. Therefore we cannot say that agents “perceive affordances” when they experience a feeling, for this would suppose either that feelings have a direct sens-

ory access to the world, which they don’t, for they extract their inputs from sensory perception, or that they have direct sensory access to the body, which they don’t have either. Feelings are the subjective counterpart of bodily changes.

Neuroscientific research about the role of emotion in perception offers evidence in favor of this view. An affordance is made immediately salient by the system’s ability to sensitively react to a (half-)perceived element in a given known context.²¹ We speak of “half-perception” on the basis of what is known about the timing of object perception. Affordance predictions are made only milliseconds after visual sensations register on the retina, i.e., before the categorisation of perceived objects is completed (Barrett & Bar 2009). The orbitofrontal cortex (OFC; involved in emotion and reward in decision making, thanks to projections from the thalamus) is able to extract an affordance in the first 80ms of the visual process, merely on the basis of low spatial frequency and magnocellular visual input (Lamme & Roelfsema 2000). What happens to perceptual access when a perceiver cannot extract affordances? Barrett & Bar (2009) have shown that the lack of emotional reactivity in early perception impairs object categorization. A patient who accidentally lost his visual ability when three years old received in adulthood a corneal transplant. In spite of his recovered ability to extract visual information from the world, this perceiver had trouble categorizing what he saw. The authors’ suggestion is that reconstituting the internal affective context associated with past exposures to an object (which was lacking in this particular case) is “one component of the prediction that helps a person see the object in the first place” (Barrett & Bar 2009, p. 1325).

In summary: the medial OFC uses early low-level visual output to match the affordance associated with it in past experience of the object: somatic markers are thereby activated, and the appropriate action is prepared. A FS enables an object to be more swiftly categorized

²¹ For a defence of this view in terms of situated cognition, see Griffiths & Scarantino (2009). The authors emphasise the environmental scaffolding that makes possible affordance detection in emoters.

at higher perceptual levels. This evidence suggests that affordances are extracted from perception, but that feelings are not themselves perceived.²² On the contrary, they offer a separate kind of feedback to cognitive perceptual processes.

4.4 Respective role of somatic markers and formal content

Let us turn now to one of the most central questions that our proposal raises. How does it explain the respective roles, in expressive intentional content, of somatic markers, on the one hand, and of the represented formal objects on the other? Cognitive theorists take emotions to represent both salient aspects of the agents' own bodily changes and an evaluative belief about an external fact, with, possibly, a causal relation between this fact and the experienced bodily change (see [Gordon 1987](#); [Tye 2008](#) and [Solomon 2007](#)). For example, when perceiving a bear in the near vicinity, one's experience is taken to be about a complex of subjective bodily impressions (a pounding heart, trembling legs, etc.) and about the perception of a bear as being the cause of these changes. Such a construal of the intentional content of feelings only makes sense within a propositional mode of thought. Can our expressive mode reflect or approximate the information contained in this complex causal structure?

Clearly, FS does not *explicitly* convey a causal relation between situation, somatic markers and subjective feeling. It carries this causal relation implicitly, however, as a consequence of the control architecture that produces feelings. In an emotional control loop, a perceived affordance causes (rather than being represented as causing) its expressive evaluation through its specialized sensory feedback. Emotional awareness expresses this functional relation. An external event (made accessible through a perceived affordance, as detailed above) is immediately followed by subjectively experienced somatic cues of a given intensity and valence. In functional terms, this sequence makes sense in

the following way. When an associated forward model has been selected (often automatically, on the basis of an environmental, somatic, or cognitive affordance), the associated sensory cues (the somatic markers in this particular episode) are automatically activated in order to monitor how this affordance is to be processed and reacted to. As has been shown elsewhere, monitoring implicitly carries information about the command (or the affordance) that is being monitored (see [Proust 2013](#)). This explanation is particularly detailed and convincing in the case of motor representations of action; the feelings of agency that result from the comparators associated with a given feedforward model express (among others) whether the emoter is, or is not, the author of the action currently attended to (see [Wolpert et al. 2001](#) and [Pacherie 2008](#)). The present proposal generalizes the functional significance of feelings throughout their diverse types (reviewed in section 2). As the outcome of sensory comparators, feelings always carry a structured information set about the type of affordance they contribute to regulating, about its amount, and about which actions are appropriate. This information, in its own expressive mode, functionally approximates a causal relation that is, when propositionally expressed, represented as a relation between an internal state, an external cause, and a disposition to act.

In summary: Feelings do not gain their aboutness through a propositional thought where the contrast between object and property is semantically marked; they gain their functional (rather than propositional) aboutness (f-aboutness) through the respective roles, in adaptive control, of the selection of an affordance-dependent control model and of the markers that allow comparisons of valence and intensity to be expressed.

4.5 The attribution problem

This account, however, fails to explain observed variability in the production of feelings and the interpretation of what feelings are “about”. There are cases where agents misattribute their sadness, their anger, or their happiness to an

²² When we say that a feeling is felt, “felt” is not intended to mean “perceived”, but, rather, “entertained”.

event that is either not real, or that actually played no role in feeling production. How can such a misattribution be explained on the present proposal? Our first attempt to address this question is based on the subjective grounding of affordances. “Feeling *f*” normally means that an affordance is sensed, expressed, and subjectively represented as present. This does not mean that the affordance has an objective counterpart. Thus a thirsty traveller can be delighted or relieved when subjected to a water mirage. It is no problem for this view, then, if an event does not have the action potential for a given affordance it is expressed as having.

A trickier problem for the proposal is that a person might feel an *f*-feeling while she thinks that she has a *g*-feeling. Is such a situation even possible? To deal with this question, we must first clarify what “transparency” means when applied to feelings. A mental state is transparent if, when it is activated, its intentional content is accessible to the subject who entertains it, while its vehicle properties are not. On the view defended above, feelings are transparent, because their somatic markers are felt in connection with a certain affordance, and because their valence and intensity directly influence the emoter’s motivation to act in a given way. Such transparency, however, does not need to entail the subject’s ability to verbally report the content of her feeling. First, as seen above, a feeling can be felt by a nonhuman or by a child, both of whom lack the requisite verbal and conceptual capacities. Second, even an agent endowed with language can express through somatic markers a feeling with a distinctive FS content while failing to accurately report, in conceptual terms, what her feeling is “about”. We saw that [aboutness], i.e., reference to an independent event or object, is not a concept that belongs to FS. When subjects try to infer [aboutness] from their experience, their propositional system of representation (PS) is solicited. Because the latter has an analytic rather than an evaluative function, additional constraints step in. While nonconceptual, intensive (analog) and value considerations and norms regulate FS, conceptual, digital, and in-

strumental considerations and norms regulate PS.²³

Hence, when having to report about her feelings, a subject needs to translate one mode of representation into another, with no guarantee that this translation will not enrich or modify FS intentional content. First, she may no longer have access to the rich diversity of her FS experience, because her attention is no longer directed toward the relevant contextually-activated affordance. Second, she has to monitor other goals and their corresponding (social, instrumental, or epistemic) norms. For example, she needs to present her feelings to herself and to others in a socially acceptable way, and to try to justify them rationally. This in turn will depend on her existing background beliefs, on her self-concept, on her capacity for making self-attributions of this particular kind, and on her willingness to perform this kind of introspective report. A number of experiments and novels have documented the wide gap between people’s feeling experiences and the verbal report they provide, or the reasons they offer, for having this or that feeling. These considerations suggest, then, that the issue of transparency cannot be adjudicated independently of one’s viewpoint about mental architecture.²⁴ According to the present proposal, an affordance is *first* subjectively recognized through the resonance it produces—through its specific feeling, rather than through a concept-based interpretation.

Let us now return to our earlier question. Can a person actually feel an *f*-like feeling, and mistake this *f*-feeling for a *g*-feeling? According

²³ About the nature and role of nonconceptual norms, see Proust (2013).

²⁴ An alternative proposal by Carruthers (2011) sees as a condition of transparency of an affective feeling, rather, that the corresponding appraisal include the detection of the details of the associated non-conceptual somatic markers, which makes the recognition of a specific emotion possible, as well as its subsequent global broadcast—hence making this information available to the mindreading system. This analytic view of feelings, however, makes it utterly mysterious how a given pattern of autonomic measures is ever recognized, among thousands of similar patterns, as distinctive of an emotion. On the present view, a feeling is produced within a given forward model, which automatically activates the comparator for this affordance. Transparency, then, is effective only when a given forward model is activated, and does not need to transfer to a verbal modality. This seems to be recognized in part by Peter Carruthers, when he concludes that “we can have transparent access to the strength of only our occurrent *context-bound* affective attitudes” (2011, p. 146).

to the present account, this situation would presuppose that an *f*-feeling, as it occurs in the expressive mode, is misdescribed in a verbal report as a *g*-feeling, to finally be genuinely felt to be *g*. On this view, a change in representational form would not only make it possible to reinterpret the initial experience in terms of a different one, but also to feel differently. To see whether this case is plausible, it is worth discussing Schachter and Singer's (1962) adrenaline experiment.

5 Do beliefs influence affective report?

Schachter and Singer's famous adrenaline study aimed to collect evidence in favor of a two-factor theory of emotion, according to which a changed state of arousal leads agents to form feelings with a given valence that depends only on the epistemic/motivational context. Participants' arousal was manipulated by injecting them, under pretext, with adrenaline or a placebo. Only a subgroup of the adrenaline participants were informed that they had received a drug that would modify their arousal level. Participants were subsequently invited to stay in a waiting room where a confederate was either pretending to be euphoric or angry. Participants' emotional responses, observed in their behavior and subsequent self-report, differed in the various conditions: those unaware of having been injected with adrenaline, and placed in the anger condition, felt angriest, followed by the placebo + anger subjects. The least angry were the adrenaline informed participants. In the euphoria condition, misinformed adrenaline participants were "somewhat" happier, adrenaline informed ones somewhat less happy (in the euphoria condition, the results failed to reach significance both for behavior and self-report).

Were Schachter and Singer successful in making the point that valence of a feeling is a matter of attribution of the source of an experienced arousal? Several powerful objections have been raised against this claim. Recall that subjects were asked to what degree they would describe themselves as happy or angry. A first problem is that the questionnaire *suggested* the relevant target categories of emotions, which is

disturbingly close to influencing participants' responses (see [Plutchik & Ax 1967](#) and [Gordon 1987](#), p. 100). Furthermore, as noted above, ex post-facto reflective labeling of one's emotion does not need to express one's original feelings. As shown by [Nisbett & Wilson \(1977\)](#), self-reporting is highly sensitive to rationalizations from context. A second problem, mentioned by the authors in the discussion, is that the subjects' verbal reports and emotional behavior failed to confirm expectations in the euphoric condition. A third methodological problem, also recognized by the authors, is that the student participants had their own independent reasons for feeling anger in passing this longish test, which predisposed them to feel anger. There are, however, more theoretical objections.

On Schachter and Singer's view, the core feeling of an emotion is an arousal change, which can be artificially induced by drugs. Valence is supposedly gained through contextual beliefs and motives. If this view is accepted, why should we expect that contextually relevant beliefs specify the feeling itself (e.g., the anger experience)? Participants may indeed have been led to believe that they were angry when they were actually merely aroused. This does not show, however, that they ever felt anything else than an arousal change ([Gordon 1987](#), pp. 100–101). Schachter and Singer may have only biased self-attributions and self-report toward target emotions. The behavioral changes that were observed and attributed to felt emotion, in addition, can be imputed to social influence, rather than to intrinsic changes.

A final worry is that inducing in a participant a somatic marker normally associated with a given feeling (e. g., increased heart rate), *and* providing the person with a context rationalizing this somatic change, does not amount to an ecological way of producing a feeling. A cognitivist theorist of emotion will insist that the mere association between a physiological cue of the feeling *f* and a context does not amount to the realization, by a participant, that she feels *f* *because* she is in such and such a context ([Gordon 1987](#), pp. 98–99).²⁵ As discussed in section

²⁵ As Gordon observes, "one will not experience fear unless one connects up that cognition with the arousal one feels. To do this re-

4, the expressive mode has a nonconceptual representation of this causal connection. The architectural relation between feelings and affordances explains why subjects experience a systematic connection between their feeling and what it is “about”, much in the same way that an agent experiences a systematic connection between an intention to move and the goal that is aimed at—that is, without needing to represent conceptually the causal connection between the two. Nothing prevents the emoter, however, from forming a secondary conceptual representation of the emotional experience she has had, and reappraising the context on the basis of her background beliefs. As a consequence of this concept-based reappraisal, the agent may either discount the relevance of her initial feeling (as in the fear-of-snake case), or redescribe it in the richer terms that she now has available (as was done, presumably, by the Schachter and Singer participants).

Taken together, these objections have led most theorists to reject Schachter and Singer’s two-factor theory of emotion, and to look for alternative accounts of the role of inferences in self-attribution of feelings. It is interesting to see, however, that a two-factor theory has also been applied to the case of M-feelings.

6 Are metacognitive feelings sensitive to beliefs and inferences?

What are metacognitive (also called *noetic*, or *epistemic*) feelings? Juxtaposing [being metacognitive] and [being a feeling] sounds, at least prima facie, dangerously close to an oxymoron. When Descartes, Locke, and other 17th-century philosophers explored the properties of ideas as being “clear”, “distinct”, “evident”, and “certain” they certainly never took them to be feelings. These notions were taken, rather, to be objective representational properties that the mind, unaided by imagination, is able to detect. David Hume, in contrast, observed in his *Treat-*

ise that “the vivacity of the idea gives pleasure”, and that “its certainty prevents uneasiness by fixing one particular idea in the mind, and keeping it from wavering in the mind of its objects” (Hume 1739/40, 2007, p. 289). Thus Hume was glad to accept that epistemic feelings exist, and that they vary in their vivacity and in their pleasantness, i.e., in their intensity and in their valence. Following Hume’s lead, let us test how our analysis of FS above fares with the case of noetic feelings. Here, again, is our proposal about the general structure of feelings.

- FS Affordance_a [Place_a=here],
[Time_a=now/soon], [Valence_a=+/-], [Intensity_{a=n}(comparatively specified on a scale 0 to 1)], [motivation to act_a of degree_d according to action program_a].

What is specific to noetic feelings is that the affordances to which the system resonates are “informational” or “metacognitive” rather than environmental. Hence, the affordance does not relate to the external environment (the “here” slot is often irrelevant, except for perceptual affordances, or place-dependent metacognitive affordances, such as concentrating in a noisy spot). Although a cognitive action does not, in general, consist in physical moves towards or away from an affordance, similar decisions are motivated or inhibited in the domain of mental agency: a high retrieval affordance motivates pursuing the memory search, a low one to quit, etc. Hence our FS analysis also applies to noetic feelings.

As already emphasized, the affordances expressed in feelings do not need to be construed conceptually in order to be detected and assessed through their associated somatic markers. A conceptual construal, however, is suggested by the names given, in the literature and in ordinary language, to M-feelings. The term “feeling of knowing” (in response, for example, to the question: “what is the capital of Australia?”) implicitly presupposes that the emoter has access to the concept of knowledge. Expressing her feeling verbally, indeed, an emoter might say: “I feel that I know the response to this question”. In this sentence, she indeed refers

quires, according to him, a second cognition: a recognition or belief that is one’s being (or taking oneself to be) in a situation of danger that is causing the arousal one feels. This “cognitivist” objection is correct when targeting S and S’s theory, who also defend a cognitivist view of feelings. The present view, however, proposes a non-doxastic account of feelings, and is thus immune to this objection.”

to her disposition to retrieve knowledge and, hence, metarepresents her knowledge disposition.²⁶ The affordance theory of noetic feelings suggests a different picture. When trying to remember a proper name, a *feeling of knowing* is a specific experience of having the ability to detect the target, and of predicting its imminent recall. It can be associated with a feeling of tension (Koriat & Levy-Sadot 1999, p. 486). This experience is associated, then, with a graded, intuitive, and affect-like appraisal of a [remembering] affordance. Rhesus monkeys working in experimental labs in comparative psychology show that they can assess their memory affordances (see Beran et al. 2012, Chapter 1).²⁷ What kind of feedback, then, do monkeys use? A surprising and substantive fact about metacognitive control, first revealed through the pioneering research of Asher Koriat, is that the comparator generating metacognitive feelings (such as a feeling of knowing in a memory task, or a feeling of clearly discriminating in a discrimination task) has no access to the semantic contents stored in memory or made available through perception. In Koriat's words, M-feelings "are mediated by the implicit application of non-analytic heuristics, relying on a variety of cues." These cues "pertain to global, structural aspects of the processing of information", such as ease of processing, time devoted to a task, familiarity, and accessibility (Koriat 2000; Koriat & Levy-Sadot 1999).²⁸ Therefore, contrary to what epistemologists have always believed, the most common type of epistemic appraisal is not directly based on the content of the thoughts to be evaluated, but on the properties of the underlying informational process.

Neuroscientific research confirms Koriat's claim. Implicit, associative cues are extracted by the working brain to select, in a cost-efficient

way, what there is to learn, to retrieve from memory, to extract from perception, or what is worth storing in memory. These are all to do with the dynamics of information processing: with its onset, with the comparative amount of activity in incompatible neural responses, and with the time needed to converge on a threshold value. Indeed, the neural activity recorded in rats' OFC when attempting to categorize olfactory stimuli was found to correlate with their predictive behavior (consisting in accepting or rejecting a task trial); similar patterns have been found in other species.²⁹

On the FS model, somatic markers have the function of expressing the intensity and valence of the noetic predictions generated from feedback at the neural level. As indicated in section 2, psychophysiological measures (electromyography) provide evidence for the existence of facial markers associated with feelings of fluency and of disfluency (Winkielman & Cacioppo 2001). Increased activity in the smile muscle, the zygomaticus major, produces feelings with a positive valence. A reduction of fluency is correlated with activity in the corrugator supercilii (involved in frowning), which suggests that this additional effort is felt as unpleasant. Intensity of positive or negative confidence, computed implicitly, is expressed by the corresponding intensity of the noetic feeling. A different somatic marker of memory appraisal is the TOT phenomenon. This often occurs when a search in memory for a specific word fails to retrieve that word within the usual time interval. The informational ingredients of FS are conveyed by the intensity of the activity in the tongue muscle, and by the affective quality of TOT. Taken together, these predict the likeli-

²⁹ See Kepecs et al. (2008). An interesting account of the predictive activity reflected in noetic feelings is that the dynamic activity in the neurons activated by a given task correlates with the so-called "accumulation of evidence" that is diagnostic of success or failure in that task. For example, in a perceptual discrimination task, where a target might be categorized as an *X* or as a *Y*, evidence for each alternative is accumulated in parallel, until the difference exceeds a threshold, which triggers the perceptual decision. The information that will generate a feeling consists, first, in the differential rate of accumulation of evidence for the two (or more) possible responses, and second, in stored information about the threshold value, computed from prior trials, which the rate of accumulation should reach in order to make a cognitive decision likely to be correct. For a discussion and review of the literature, see Fleming & Dolan (2012), and Proust (2013, pp. 99).

²⁶ Arango-Muñoz (2012) claims that feelings of forgetting and feelings of knowing are cases of "conceptual experiences". According to the present view, following the lead of Koriat and colleagues, M-feelings can overlap with judgments, and be redescribed in conceptual terms; they pertain, however, to different representational levels. There are no "conceptual experiences", except in the sense of experiencing the comparative fluency of concepts.

²⁷ As indicated above, rhesus monkeys are able, in a perceptual or memory task, to opt out of more or less challenging trials as a result of trial difficulty.

²⁸ As will transpire below, all these cues are, as far as we know, dimensions or effects of fluency, i.e., of ease of processing.

hood of successful retrieval. An implicit cue-based heuristic might thus explain why TOTs have the valid predictive value they do (Schwartz et al. 2000).

6.1 Two-factor theories of M-feelings

In our FS single-factor model, M-feelings have an intrinsic intensity and an intrinsic valence. Two-factor theories make a different claim, in ways analogous to Schachter and Singer's theory of aboutness in affects: M-feelings have an intrinsic arousal level, but their valence depends on the environment. Jacoby and his colleagues were the first to embrace a two-factor view about feelings of fluency. They manipulated participant's exposure to an item in order to show that enhanced fluency generates an illusory feeling of familiarity. Under conditions of divided attention, reading a list containing both famous and not famous names raised participants' disposition to wrongly judge as famous some names presented in a second list, merely because these names had already been read in the first list. Schachter and Singer's idea was that fluency is a generic feeling, that needs to be interpreted on the basis of goals and current cues, in order to deliver a qualitatively different specific feeling:

Inherent in the idea that the subjective experience of familiarity arises from an interpretation of cues is the notion that cues can be interpreted in a variety of ways. As noted above, if ease of identifying an item is obviously being manipulated by the experimenter, the resulting perceptual fluency does not give rise to a feeling of familiarity. Attributions are also affected by one's goals. In the context of attempts to remember, people may be more likely to interpret ease of generating an item or perceiving it as familiarity. In the context of other tasks, the same cues may be interpreted in other ways. (Kelley & Jacoby 1998, p. 129)

From their viewpoint, the fluency generated by a given name can, according to the task and the

information made consciously available to a participant, be experienced as a feeling of familiarity, or as a feeling of recognition of that name as "old" (i.e., presented in a former list). They conclude that a feeling of fluency (generated by a perceived name) will be experienced as a function of the alternative ways of interpreting this feeling, on the basis of the agent's goals and the additional cues available.³⁰

A similar two-factor theory has been defended in the (Whittlesea & Williams 2000; Whittlesea & Williams 2001) model of M-feelings. According to this model, feelings of familiarity result from the perception of a *nonspecific* discrepancy between the expected and the observed rate of processing of elements in a given context. Valence and the associated action guidance, on the other hand, are based on a conceptual interpretation of what this discrepancy means. For example, you find yourself waiting for the bus next to people you expect to be total strangers. Suddenly, you have an unexpectedly high fluency experience when looking at the face of someone you have already encountered several times—a clerk from the local grocery shop. This unexpectedly high rate of discrepancy-reduction determines an intense feeling of familiarity with a strong motivation to identify the familiar face (see Whittlesea & Williams 2001). Had you seen the clerk in the local grocery store instead, you would have merely had a feeling of recognition when seeing the clerk.

To summarize: the core idea in two-factor accounts is that participants have a primary feeling of fluency, which they interpret in more specific terms as a function of their goals and of the context as they consciously represent it to be. Thus, on this view, a feeling partly relies on background knowledge, and partly on a naïve theory concerning the relation between feelings and mental activity (Schwarz & Clore 2007). The naïve theory is as follows: feelings are about what one is doing, so this feeling must be about this event of trying to perceive, or this attempt at retrieving, etc.

³⁰ Jacoby & Whitehouse (1989) similarly argue that a feeling of fluency can be experienced as familiarity in a memory task, and as confidence in a problem-solving task.

As already observed above, a naïve-theory view is incompatible with monkeys' and young children's epistemic evaluations based on fluency. Our FS structure offers an alternative account: cues (associative heuristics) dictate how an affordance is detected, assessed, and exploited in a context, but these cues are not consciously available, and hence do not depend on a naïve theory of the task. The Jacoby and Whitehouse evidence is compatible with a procedural view of engagement in a task through automatic memory processes, and of the feelings of familiarity they generate. A comparator is always activated as a function of a subject having been highly trained in the corresponding first-level cognitive task. Monkeys and humans feel that a memorial or perceptual affordance is present because, if they need to assess whether, for example, an item was seen earlier, the associated comparator produces a feeling of a given intensity and valence indexing the remembering affordance. Thus, it is uncontroversial that a context-dependent factor determines both the task to be performed and the reactive metacognitive feeling about this task.

It does not follow from the context-dependence of a cognitive task, however, that a concept-based interpretation will affect the experienced feeling itself, as maintained by the two-factor theorist. A cue-based, non-analytic heuristic is not inferential in the interpretive, first-person sense. Regrettably, the word "inference" has been loosely used in affective and in metacognitive studies, to refer both to "automatic, non-analytic, largely unconscious and fast associative processes" (Nussinson & Koriat 2008) and to conscious reasoning and theory-building (Schwarz & Clore 2007). These two types of processes (respectively called "automatic" and "controlled"), are now held by many authors to operate independently.³¹ While unconscious heuristics rely on implicit associations between cues, inferences comprise deductions from premises to conclusions. Looking back at Jacoby and Kelley's point above, we see that

the authors are referring to unconscious cues being recruited for a task: they are thus referring to unconscious associative heuristics rather than to explicit concept-based reasoning. The memory interactions they are exploring, however, typically involve both automatic and controlled processes, which is a source of confusion. As Jacoby and Kelley are eager to show, implicit associations and explicit reasoning lead to different, incompatible predictions. As a result, the evidence they present shows how automatically-generated feelings can be theorized about in controlled processes. It does not demonstrate, however, that feelings depend upon theorization. A theory of the task, in contrast with automatically generated feelings, offers reasons to attribute to oneself beliefs and motivations to act, and, possibly, to reject the relevance of feelings for any particular task.

Our proposal, then, has several advantages over inferential or theory-based accounts of fluency. First, it explains why a feeling of fluency can be experienced, and why it can motivate agents' metacognitive responses in species or individuals with no concept-based attributive capacity (i.e., with no capacity for mindreading). Second, our proposal accounts for the difference between a type of M-feeling (a feeling of fluency) and the various ways in which it is experienced across cognitive tasks. Granting that comparative ease of processing can *always* be computed, and can be used as a reliable indicator of the likelihood of success across a wide range of cognitive activities, it is not surprising that there is a type of feeling based upon it. Fluency can be perceptual, memorial ("retrieval fluency"), or conceptual. It can be used in predictive or retrospective evaluations. If agents are asked to determine which statements are likely to be true or false (presumably a question that only—but not all—humans can understand), felt perceptual fluency will induce a "truth effect". Agents will evaluate a statement as more likely to be true than another merely because it is easier to read.³² If agents are asked

³¹ For a defence of the distinction see Jacoby & Brooks (1984), Koriat & Levy-Sadot (1999), Recanati (2002) and Smith & DeCoster (1999). Koriat & Levy-Sadot (1999) both emphasize the distinction and use the term "inference" in both cases.

³² There is abundant evidence, however, that M-feelings uncritically guide epistemic decision (i.e., are unopposed by concept-based processes) mostly when the cognitive task is unimportant, when cognitive resources are limited (under time pressure or divided attention), and when agents are in a good mood (Nussinson & Koriat 2008; Schwarz 2004).

to detect faces of known people (or of stimuli previously shown), felt fluency will generate a sense of familiarity, which motivates agents to try to identify the target. If people are asked to assess the frequency of a given phenomenon, felt retrieval fluency—that is, what comes immediately to mind—will be used to judge what is more frequent. Felt fluency will also have effects outside of metacognition: if agents are asked which particular face, landscape, or picture they prefer, felt fluency will influence their decision. Several affordances, then, may be associated with the same globally expressive *type* of feeling (constructed as the set of feelings with the same type of facial markers for ease of processing, for example). The notion of type of feeling is a technical term, which is useful to distinguish the diverse ways in which fluency is used by the brain. But a type of feeling is never experienced; only tokens of the type are. Tokens of feelings of the same type will differ in the specific affordances that are detected, and in the tendencies to act that the feeling motivates. As a consequence, one cannot say that feelings of fluency “feel the same” to an emoter: fluency experienced in an FOK and in an FOR, for example, apply to different segments of processing, assess different things, and motivate a different action program. You may first have an FOK after a question is addressed to you, and then fail to have the associated FOR after having come up with a response. These differences have nothing to do with an interpretation: they are constitutive of what sensitivity to a given affordance amounts to. Take the case of feelings of familiarity. As summarized [above](#), Whittlesea and Williams claim that fluency is the core of the experience, while familiarity is a conceptual interpretation of this core feeling. It is more economical, however, to suppose that familiarity is a different feeling within the general fluency type, and that it is associated with a different affordance.

In summary: engaging in a particular cognitive task (e.g., trying to remember, evaluating retrieval, assessing frequency) does not need, per se, to involve a naïve theory of the task. It only requires having a salient affordance, and an implicit heuristic for metacognitive predictions in that task.

6.2 Incidental versus integral feelings

Our proposal also allows us to address in affective terms the issue of incidental versus integral feelings, which, in the literature, is invariably framed in inferential terms (with all the ambiguity relating to this expression). Metacognitive feelings are called “incidental” when they are not based on valid cues for the cognitive task at hand, and hence, have no predictive value. They are called “integral” when they actually carry information about cognitive outcome. Granting the universal role of fluency in metacognition, how do people know when a feeling of fluency is relevant to a given task, and which sequence of their cognitive activity needs to be monitored? A frequent answer, in the literature, is that agents believe that fluency applies by default to the present domain of judgment. When, however, agents are led to believe that a feeling of fluency is purely incidental to the task at hand, they will discount it in their decision, on the basis of a theory of the domain of interest (see [Schwarz & Clore 2007](#) and [Whittlesea & Williams 2000, 2001](#)). Let us suppose, in what we shall call case (a), that an agent is explicitly told that a given cue, such as the ease of reading a given sentence, is irrelevant to a given task—such as assessing the truth value of the written statement. Or, alternatively, let us suppose—case (b)—that the agent discovers by himself that there is a connection, but with reverse relevance. Perhaps he finds that badly written sentences, involving added processing effort—in a given context—are likely to be true (see [Unkelbach 2007](#) and [Unkelbach & Greifeneder 2013](#)). A popular account of these cases is that people will infer respectively, for (a): that the feeling of fluent reading they have had *is not about* the target task, which entails that reading fluency does not predict truth, or, for (b): that what predicts the truth of a written utterance, in this particular context, is disfluent reading (see [Schwarz & Clore 2007](#), p. 394).

According to this two-factor account, M-feelings are cognitively penetrable. They can be suppressed at will, on the basis of a reinterpretation of their being experienced, or can even be

used to predict falsity instead of truth.³³ On the account proposed here, in contrast, M-feelings are never cognitively penetrable. Why, then, do subjects stop trusting their feeling of fluency? Our answer is the following. In the first type of case, subjects do not allow their feelings of fluency to guide their decision because they have received verbal instructions to this effect. In the second type of case, subjects no longer use their feelings of fluency to form an epistemic decision in the proposed task, because they have learned, over time, that these feelings do not predict truth in this task.

In case (a), then, subjects are confronted with a different task. They are no longer asked to express their confidence in the truth of a given sentence (an intuitive, associative task); they are asked to assess the truth of sentences by taking into account the fact that their feelings of fluency are irrelevant. This new task requires the participants to form appraisals based on analytic reasoning. Feelings no longer drive their evaluation and epistemic decision.

In case (b), where bad writing is associated with likely truth, no “theory of the task” needs to be formed, on top of the first-order task, which consists in judging whether a written statement is true or not. A mere change in cue validity can produce, over time, a change in associative heuristics, and, hence, in feelings and in decisions to act. For example, just as our thirsty traveller will eventually learn not to trust an apparent “drinking affordance”, an agent will learn, in certain recurrent contexts, not to trust an apparent “fluency affordance”. Obviously, cue validity can, in humans, be conveyed verbally; this will considerably abridge the revision process of the associated program of action. We then return to case (a): participants will be able to immediately discount an apparently valid cue, to turn to analytic appraisals, and to refrain from acting on their fluent feeling (which, however, is still there). Cue validity, however, can be learnt implicitly as

well, which weakens the case for a theory-laden view of feelings.

These observations suggest that feeling-based and analytic appraisal, as hypothesized in this proposal, “tap separate databases representing knowledge in different formats.”³⁴ A feeling of fluency, as a result, can survive being discounted in decision-making. Another finding points in the same direction. There is evidence that, even when an M-feeling has been explicitly discounted (i.e., shown to agents to unduly bias their epistemic assessment), the initial feeling remains unaffected, and is able to promote further epistemic decisions. In Nussinson & Koriat’s (2008) study, agents exposed to unsolved anagrams and to anagrams accompanied by their solution, were asked to rate the difficulty of these anagrams for naïve participants with no prior access to the solution. The participants’ ratings were influenced by the differential fluency that the anagrams presented for them: the higher fluency of solved anagrams biased their attributions of difficulty. After being informed of the contaminating effect of knowing the solutions, the participants were invited to correct their attributions by re-rating the difficulty of the anagrams, which they did. However, the participants were subjected to a subsequent test, where, under time pressure, they had to predict which of two anagrams would be harder for others to solve. These other-attributions of difficulty presented, again, the same bias for known anagrams. Being under pressure allowed participants’ M-feelings to guide decision. The verbal instruction could shift their controlled responses when re-rating the anagrams, but did not lead the participants to recompute them, as should have been the case if feelings are cognitively penetrable.

In summary: what participants learned (that solved anagrams only *look* easier to process) did not influence what they felt later (higher fluency is diagnostic of ease of solving).

7 Are all feelings affective?

It is often noticed that a phenomenological contrast seems to exist between feelings—that is, they are not equally emotional. Are not M-feel-

³³ This two-factor account is endorsed by Unkelbach (2007): “the feeling resulting from the discrepancy is non specific, and the discrepancy triggers a search for an explanation [...]. The experienced variations are not attributed to prior exposure, resulting in a feeling of familiarity, but to some other quality of the statement, namely, that a statement is true.”

³⁴ A quote from Smith & DeCoster (1999), p. 329, who offer a strong defence of this view.

ings in general as “cold” as the proprioceptive feeling that my right arm is being extended? Or can they also be “hot”—that is, involve valence, i.e., be pleasant or unpleasant? Our proposal of a common expressive evaluative format suggests that all the feelings vary in affect in roughly the same way, because they all include valence in their informational structure. [Stepper & Strack \(1993\)](#), however, have emphasized that epistemic feelings are “cold”. Feelings like effort, familiarity, surprise, or feeling of knowing “have no fixed valence”, in the sense that they don’t feel particularly good or bad. Linguistic research on the emotional lexicon is invoked as congruent evidence: for words referring to readiness, success, and a desire to deal with new information (like “alert” “confused”), i.e., terms expressing metacognition, affects are not “focal”, which implies that they are not centrally emotional ([Ortony et al. 1987](#)).

There is abundant evidence, however, that feelings of fluency increase perceivers’ liking of the objects perceived. Familiar items (other things being equal) are found to be more pleasant than new ones. An initially neutral stimulus is felt to be pleasant after repeated exposure. This “exposure effect”, first demonstrated by Zajonc, has been attributed to increased perceptual fluency ([Zajonc 1968](#)). This affective effect of fluency has since been found to apply to any dimension of a perceptual input. The sense of beauty in a symmetrical face or in a landscape, or the pleasure felt in contemplating a picture seem to be inherent to the feeling of fluency generated in the perception. As noted above, psychophysiological measures in the facial muscles provide additional evidence for the affective character of the feeling of fluency ([Reber et al. 2004](#); [Winkielman & Cacioppo 2001](#); for a review see [Oppenheimer 2008](#)).

An interesting, untested, speculation intended to explain the presence of cold and hot versions of feelings is that valence, although never fully absent from monitoring, is modulated by dynamic aspects of the task under evaluation ([Carver & Scheier 1990](#); [Carver & Scheier 2001](#)). On this view, affective feelings can appear in physical and cognitive action, and probably also in somatosensory experience,

when certain dynamic conditions for affective reactions are present. But what are these conditions?

Let us first examine an area where these dynamic conditions seem to have a minimal role. This is the area of first-order motor control (including the initiation of an action, the monitoring of its development, and of goal completion). As with any other form of control, motor control involves specialized feelings, in the above sense of subjective experiences with a distinctive embodied phenomenal quality (see [Pacherie 2008](#)). At first glance, these feelings do not typically seem to be affective.³⁵ Why is this so? According to Carver and Scheier, this can be explained by the dynamics of a monitored activity that generates feelings. Affective feelings are part of a second-order type of feedback, having, in their terms, “the meta-monitoring function” of “checking on how well the action loop is doing at reducing the behavioral discrepancy that the action loop is monitoring”. This meta-loop, then, monitors a particular aspect of one’s progress in relation to one’s distal goal: it represents “*the rate of discrepancy reduction in the behavioral (monitoring) system over time*”. This dynamic representation is what a feeling is equipped to offer: the intensity and quality of a positive, or a negative, feeling express how far above, or how far below, the observed *rate* of discrepancy reduction is, with respect to some reference value. One consequence of this view, if it turns out to be experimentally validated, is fascinating and deep: affect in action does not depend merely on the amount of discrepancy being reduced. An agent may be an inexperienced performer in a task; if the velocity of her progress to the goal is higher than expected, she will feel more confident, and have retrospectively more positive feelings when reaching her goal than a competent performer whose progress to the goal is as steady as predicted.

There is a second type of affect, according to Carver and Scheier, that the dynamics of prediction can generate. Acceleration is the rate of change of velocity. Feelings express such acceleration when the rate of discrepancy reduc-

³⁵ Even in this domain, however, an error signal, when conscious, is associated with an unpleasant feeling.

tion *increases* beyond expectancy—a sense of exhilaration then occurs. Lucky athletes, who break several records within days, experience this. Symmetrical feelings of sinking, or despair, arise when the rate of discrepancy reduction *decelerates* unexpectedly and falls below the expected threshold more quickly than anticipated. In summary, cold motor feelings are generated when one is routinely acting on the world, when things develop as expected, except for small motor adjustments. Hot action feelings are generated when action monitoring involves unexpected dynamics of reduction or increment of likely success or failure.

How does this theory apply to M-feelings? A similar contrast may exist in M-feelings. Carver and Scheier's model allows us to predict that M-feelings can have colder and warmer varieties, depending on the dynamics of the discrepancy reduction that they express. As seen above, there are two varieties of M-feelings, distinguished by their function. Some, like FOKs, have a predictive function. Others, like FORs, perform retrospective evaluation. Neuroscientists explain these feelings through the rate of the accumulation of evidence, measured through the comparative activity of the neural assemblies involved in cognitive decision. (This rate of accumulation has to be compared with a stored standard in order to produce a reliable feeling of confidence.) From this widely accepted model, it follows that the rate of reduction of discrepancy toward a confidence threshold is automatically computed, and plausibly expressed through somatic markers that themselves have a varying intensity.

If this reasoning is correct, then although all M-feelings do not often have a definite “hot” quality comparable to fear and love, they always have a valence, according to whether they predict an agent's progress towards or away from her cognitive goal. To find more intense M-feelings, however, one needs to look at the dynamics of *meta-monitoring*, which is when an agent expects a given rate of reduction of the discrepancies toward her cognitive goal, and either observes a rate that is well above the expected rate or well below it. In these cases, the sense of confidence that the positively surprised agent

experiences is modulated by an intense, highly motivating affect of joy and renewed passion for the associated cognitive activity; while the uncertainty of the negatively surprised agent is associated with an intense, highly demotivating affect of discouragement, or loss of interest. Note how crucial an intense feeling of this kind can be, especially with regard to future motivation. It can precipitate in children a passion for learning; or it can lead them to reject an activity, or even a whole group of similar activities, because of the threatening affect associated with the activity, often combined with a still more threatening social affect (the sense of being an inferior, incompetent performer, or of being stupid). This kind of meta-monitoring cognitive affect, important as it is in predicting and fuelling epistemic motivation, is not easily observable in experimental settings, because it is elicited in middle or long-term forms of cognitive tasks, such as studying at school in a given grade, learning algebra, etc. This may in part explain why Stepper and Strack have failed to encounter it.

To summarize: noetic feelings, like all feelings, have an evaluative function. They are the output of a monitoring process, which expresses how likely it is that an agent's cognitive preferences or goals will be (or have been) fulfilled in a given task and context. They all have a valence, but their affective tonality is more intensely felt in special cases that arise when meta-monitoring makes “intensively new” affordances salient. The rate or the acceleration with which an observed initial discrepancy differs from a predicted standard value may either exceed the expected value, thereby producing positive feelings of confidence or feelings of knowing, or be insufficient to reach this value, producing negative feelings of uncertainty. The intensity of positive or negative affect in M-feelings thus depends on particularly unexpected properties of the underlying cognitive activity.

8 Conclusion

On the present proposal, “feelings” are not isolated sensory events. They are, rather, the

ingredients of a nonlinguistic expressive mode that allows organisms to evaluate and predict environmental changes and affordances. This expressive mode is of a relational, intensive kind that is not suitable for a predicative, concept-based representation of the world. As a consequence, feelings are not themselves judgments about the world or about one's own thoughts. They are not "about" anything in the objective, referring sense of the term. Feelings are able to approximate (in their own mode) the guidance offered by full-blown judgments, and hence can be re-described in conceptual terms when the latter are available to the emoter.

The importance of the duality between an expressive and a propositional system of representation has generally been overlooked. Even dual-processing theorists rarely appreciate that the two systems involved in cognitive evaluation and in reasoning have their own independent, although asymmetrical, role to play. A purely automatic, reactive type of evaluation is possible, and is present in nonhumans and young children. It is prone, however, to generating throughout life illusions of competence and reasoning errors. A conceptually-controlled type of evaluation, on the other hand, can partially inhibit the influence of the expressive system, but it still depends on the latter to weigh the impact of context on ability, and to assess the trade-off between ease of processing and informativeness—that is, relevance—that is crucial in communication and in problem solving.

A major practical consequence of the duality between the two target representational modes concerns pedagogy. Children cannot learn what they are *not* motivated to learn. Their motivation heavily depends on their subjective experience of what a school context affords them. Their feelings of confidence, i.e., the feedback from the cognitive tasks they engage in, have to be sufficiently positive and appropriately calibrated in order for them to form their own realistic and motivating cognitive goals. No amount of analytic reasoning can replace a positive experience when learning.

Acknowledgement

I am grateful to Dick Carter, Laurence Conty, Terry Eskenazy, Martin Fortier, Jonathan Frome, Thomas Metzinger and Jennifer Windt for their critical comments. Special thanks to Dick Carter for his linguistic advice, to Tony Marcel, whose critical questions about the commonality between M-feelings and affective feelings inspired the present article, and to Robert Gordon, for giving me access to some of his unpublished writings. This research has been supported by an ERC Senior Grant "Dividnorm" #269616, and by two institutional grants: ANR-11-LABX-0087 IEC and ANR-11-IDEX-0001-02 PSL.

References

- Arango-Muñoz, S. (2012). The nature of epistemic feelings. *Philosophical Psychology*, 27 (2), 193-211. [10.1080/09515089.2012.732002](https://doi.org/10.1080/09515089.2012.732002)
- Barrett, L. F. & Bar, M. (2009). See it with feeling: Affective predictions during object perception. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364 (1521), 1325-1334. [10.1098/rstb.2008.0312](https://doi.org/10.1098/rstb.2008.0312).
- Bechara, A., Damasio, H. & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, 10 (3), 295-307. [10.1093/cercor/10.3.295](https://doi.org/10.1093/cercor/10.3.295).
- Beran, M. J., Brandl, J., Perner, J. & Proust, J. (Eds.) (2012). *The foundations of metacognition*. Oxford, UK: Oxford University Press.
- Botvinick, M. & Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature*, 391 (6669), 756-756. [10.1038/35784](https://doi.org/10.1038/35784).
- Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin*, 109 (2), 204-223. [10.1037/0033-2909.109.2.204](https://doi.org/10.1037/0033-2909.109.2.204)
- Carruthers, P. (2011). *The opacity of mind. An integrative theory of self-knowledge*. Oxford, UK: Oxford University Press.
- Carver, C. S. & Scheier, M. F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review*, 97 (1), 19-35. [10.1037/0033-295X.97.1.19](https://doi.org/10.1037/0033-295X.97.1.19)
- (2001). *On the self-regulation of behavior*. Cambridge, UK: Cambridge University Press.
- Damasio, A. (1994). *Descartes' error: Emotion, reason and the human brain*. New York, NY: Putnam.
- (2003). *Looking for Spinoza: Joy, sorrow and the feeling brain*. Orlando, FL: Harvest Book Harcourt, Inc.
- Decety, J. & Meyer, M. (2008). From emotion resonance to empathic understanding: A social developmental neuroscience account. *Development and Psychopathology*, 20 (04), 1053-1080. [10.1017/S0954579408000503](https://doi.org/10.1017/S0954579408000503)
- Dezecache, G., Conty, L., Chadwick, M., Philip, L., Soussignan, R., Sperber, D. & Grèzes, J. (2013). Evidence for unintentional emotional contagion beyond dyads. *PLoS One*, 8 (6), e67371-e67371. [10.1371/journal.pone.0067371](https://doi.org/10.1371/journal.pone.0067371)
- Dolan, R. J. (2002). Emotion, cognition, and behavior. *Science*, 298 (5596), 1191-1194. [10.1126/science.1076358](https://doi.org/10.1126/science.1076358)
- Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, 6 (3-4), 45-60. [10.1080/02699939208411068](https://doi.org/10.1080/02699939208411068)
- Fleming, S. M. & Dolan, R. J. (2012). The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367 (1594), 1338-1349. [10.1098/rstb.2011.0417](https://doi.org/10.1098/rstb.2011.0417)
- Fodor, J. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.
- Frank, R. H. (1988). *Passions within reasons: The strategic role of the emotions*. New York, NY: Norton.
- Goldie, P. (2002). *The emotions: A philosophical exploration*. Oxford, UK: Oxford University Press.
- (2009). Getting feelings into emotional experience in the right way. *Emotion Review*, 1 (3), 232-239. [10.1177/1754073909103591](https://doi.org/10.1177/1754073909103591)
- Gordon, R. (1987). *The structure of emotions*. Cambridge, UK: Cambridge University Press.
- Griffiths, P. E. (1997). *What emotions really are*. Chicago, IL: The University of Chicago Press.
- (2004). Emotions as natural and normative kinds. *Philosophy of Science*, 71 (5), 901-911. [10.1086/425944](https://doi.org/10.1086/425944)
- Griffiths, P. E. & Scarantino, A. (2009). Emotions in the wild: The situated perspective on emotion. In M. Aydede & P. Robbins (Eds.) *The Cambridge handbook of situated cognition* (pp. 437-453). New York, NY: Cambridge University Press.
- Haggard, P., Clark, S. & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, 5, 282-285. [10.1038/nn827](https://doi.org/10.1038/nn827)
- Hume, D. (2007). *Treatise of human nature*. Oxford, UK: Clarendon Press.
- Jackson, P. L., Meltzoff, A. N. & Decety, J. (2005). How do we perceive the pain of others? A window into the neural processes involved in empathy. *NeuroImage*, 24 (3), 771-779. [10.1016/j.neuroimage.2004.09.006](https://doi.org/10.1016/j.neuroimage.2004.09.006)
- Jacoby, L. L. & Brooks, L. R. (1984). Nonanalytic cognition: Memory, perception, and concept learning. *The Psychology of Learning and Motivation*, 18, 1-47. [10.1016/S0079-7421\(08\)60358-8](https://doi.org/10.1016/S0079-7421(08)60358-8)
- Jacoby, L. L. & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General*, 118 (2), 126-135. [10.1037/0096-3445.118.2.126](https://doi.org/10.1037/0096-3445.118.2.126)
- James, W. (1884/1890). *The principles of psychology*. New York, NY: Dover.
- Kelley, C. M. & Jacoby, L. L. (1998). Subjective reports and process dissociation: Fluency, knowing, and feeling. *Acta Psychologica*, 98 (2), 127-140. [10.1016/S0001-6918\(97\)00039-5](https://doi.org/10.1016/S0001-6918(97)00039-5)

- Kenny, A. (1963). *Action, emotion and will*. London, UK: Routledge & Kegan Paul.
- Kepecs, A., Naoshige, U., Zariwala, H. & Mainen, Z. F. (2008). Neural correlates, computation and behavioural impact of decision confidence. *Nature*, 455, 227-231. [10.1038/nature07200](#)
- Koriat, A. (2000). The feeling of knowing: Some metatheoretical implications for consciousness and control. *Consciousness and Cognition*, 9 (2), 149-171. [10.1006/ccog.2000.0433](#)
- Koriat, A. & Ackerman, R. (2010). Metacognition and mindreading: Judgments of learning for self and other during self-paced study. *Consciousness and Cognition*, 19 (1), 251-264. [10.1016/j.concog.2009.12.010](#)
- Koriat, A. & Levy-Sadot, R. (1999). Processes underlying metacognitive judgments: Information-based and experience-based monitoring of one's own knowledge. In S. Chaiken & Y. Trope (Eds.) *Dual-process theories in social psychology* (pp. 483-502). London, UK: The Guilford Press.
- (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27 (1), 34-34. [10.1037/0278-7393.27.1.34](#)
- Lamme, V. A. & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, 23 (11), 571-579. [10.1016/S0166-2236\(00\)01657-X](#)
- LeDoux, J. E. (1996). *The emotional brain*. New York, NY: Simon & Schuster.
- Logan, G. D. & Crump, M. J. (2010). Cognitive illusions of authorship reveal hierarchical error detection in skilled typists. *Science*, 330 (6004), 683-686. [10.1126/science.1190483](#)
- Mangan, B. (1993). Taking phenomenology seriously: The “fringe” and its implications for cognitive research. *Consciousness and Cognition*, 2 (2), 89-108. [10.1006/ccog.1993.1008](#)
- (2000). What feeling is the “feeling of knowing”? *Consciousness and Cognition*, 9 (4), 516-537. [10.1006/ccog.2000.0488](#)
- Murata, A., Moser, J. S. & Kitayama, S. (2013). Culture shapes electrocortical responses during emotion suppression. *Social Cognitive and Affective Neuroscience*, 8 (5), 595-601. [10.1093/scan/nss036](#)
- Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P. & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task. *Psychophysiology*, 38 (5), 752-760. [10.1111/1469-8986.3850752](#)
- Nisbett, R. E. & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological review*, 84 (3), 231-259.
- Nussinson, R. & Koriat, A. (2008). Correcting experience-based judgments: The perseverance of subjective experience in the face of the correction of judgment. *Metacognition and Learning*, 3 (2), 159-174. [10.1007/s11409-008-90](#)
- Oppenheimer, D. M. (2008). The secret life of fluency. *Trends in Cognitive Sciences*, 12 (6), 237-241. [10.1016/j.tics.2008.02.014](#)
- Ortony, A., Clore, G. L. & Foss, M. A. (1987). The referential structure of the affective lexicon. *Cognitive Science*, 11 (3), 341-364. [10.1207/s15516709cog1103_4](#)
- Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition*, 107 (1), 179-217. [10.1016/j.cognition.2007.09.003](#)
- Plutchik, R. & Ax, A. F. (1967). A critique of determinants of emotional state by Schachter and Singer (1962). *Psychophysiology*, 4 (1), 79-82. [10.1111/j.1469-8986.1967.tb02740.x](#)
- Prinz, J. (2004). *Gut reactions. A perceptual theory of emotions*. Oxford, UK: Oxford University Press.
- Proust, J. (2000). Awareness of agency: Three levels of analysis. In T. Metzinger (Ed.) *The neural correlates of consciousness* (pp. 307-324). Cambridge, MA: MIT Press.
- (2009). The representational basis of brute metacognition: A proposal. In R. Lurz (Ed.) *Philosophy of animal minds: New essays on animal thought and consciousness* (pp. 165-183). Cambridge, UK: Cambridge University Press.
- (2013). *The philosophy of metacognition. Mental agency and self-awareness*. Oxford, UK: Oxford University Press.
- Reber, R., Fazendeiro, T. A. & Winkielman, P. (2002). Processing fluency as the source of experiences at the fringe of consciousness. *Psyche*, 8 (10), 1-21. [10.1155/6152](#)
- Reber, R., Schwarz, N. & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, 8 (4), 364-382. [10.1207/s15327957pspr0804_3](#)
- Recanati, F. (2002). Does linguistic communication rest on inference? *Mind & Language*, 17 (1-2), 105-126. [10.1111/1468-0017.00191](#)
- Schwartz, B. L., Travis, D. M., Castro, A. M. & Smith, S. M. (2000). The phenomenology of real and illusory tip-of-the-tongue states. *Memory & Cognition*, 28 (1), 18-27. [10.3758/BF03211571](#)

- Schwarz, N. (2004). Metacognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology*, 14 (4), 332-348. [10.1207/s15327663jcp1404_2](https://doi.org/10.1207/s15327663jcp1404_2)
- Schwarz, N. & Clore, G. L. (2007). Feelings and phenomenal experiences. In A. W. Kruglanski & E. T. Higgins (Eds.) *Social psychology: Handbook of basic principles* (pp. 385-407). New York, NY: Guilford.
- Smith, E. R. & De Coster, J. (1999). Associative and rule based processing. In S. Chaiken & Y. Trope (Eds.) *Dual-process theories in social psychology* (pp. 323-336). New York, NY: Guilford.
- Solomon, R. C. (2007). *True to our feelings: What our emotions are really telling us*. New York, NY: Oxford University Press.
- Stepper, S. & Strack, F. (1993). Proprioceptive determinants of emotional and nonemotional feelings. *Journal of Personality and Social Psychology*, 64 (2), 211-220. [10.1037/0022-3514.64.2.211](https://doi.org/10.1037/0022-3514.64.2.211)
- Strawson, P. F. (1959). *Individuals*. London, UK: Methuen.
- Thompson, V. A., Prowse Turner, J. A. & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology*, 63 (3), 107-140. [10.1016/j.cogpsych.2011.06.001](https://doi.org/10.1016/j.cogpsych.2011.06.001)
- Tye, M. (2008). The experience of emotion: An intentionalist theory. *Revue Internationale de Philosophie*, 243 (1), 25-50.
- Unkelbach, C. (2007). Reversing the truth effect: Learning the interpretation of processing fluency in judgments of truth. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33 (1), 219-230.
- Unkelbach, C. & Greifeneder, R. (2013). A general model of fluency effects in judgment and decision making. In C. Unkelbach & R. Greifeneder (Eds.) *The experience of thinking: How the fluency of mental processes influences cognition and behavior* (pp. 11-32). London, UK: Psychology Press.
- Whittlesea, B. W. & Williams, L. D. (2000). The source of feelings of familiarity: The discrepancy-attribution hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26 (3), 547-547. [10.1037/0278-7393.26.3.547](https://doi.org/10.1037/0278-7393.26.3.547)
- (2001). The discrepancy-attribution hypothesis: I. The heuristic basis of feelings and familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27 (1), 3-13. [10.1037/0278-7393.27.1.3](https://doi.org/10.1037/0278-7393.27.1.3)
- Winkielman, P. & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation elicits positive affect. *Journal of Personality and Social Psychology*, 81 (6), 989-1000. [10.1037/0022-3514.81.6.989](https://doi.org/10.1037/0022-3514.81.6.989)
- Wolpert, D. M., Ghahramani, Z. & Flanagan, J. R. (2001). Perspectives and problems in motor learning. *Trends in Cognitive Sciences*, 5 (11), 487-494. [10.1016/S1364-6613\(00\)01773-3](https://doi.org/10.1016/S1364-6613(00)01773-3)
- Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, 9 (2 pt 2), 1-27. [10.1037/h0025848](https://doi.org/10.1037/h0025848)
- (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35 (2), 151. [10.1037/0003-066X.35.2.151](https://doi.org/10.1037/0003-066X.35.2.151)

The Extension of the Indicator-Function of Feelings

A Commentary on Joëlle Proust

Iuliia Pliushch

In the following commentary I will first briefly review the target article, then voice some critical points, and last offer a positive proposal according to which tension in self-deception is a kind of a metacognitive feeling. Proust offers a novel, inspiring view that feelings possess an indexical (non-conceptual) format, are transparent (that is, they may be re-described in propositional terms, but not thereby changed), and acquire valence if the rate of change towards fulfilling the given affordance is greater or less than expected. In my critique I will first point to difficulties in disentangling feelings from emotions, then try to provide a more precise description of the formal object of feelings, along with some examples, and offer a definition of “directness” that is consistent with predictive coding—as well as argue that feelings might be influenced by concepts even if they themselves are non-conceptual. Last, I propose that tension in self-deception is a metacognitive feeling.

Keywords

Affective feelings | Appraisal | Metacognitive or noetic feelings | Predictive coding | Self-deception | Two-factor account

Commentator

[Iuliia Pliushch](#)

pliushi@students.uni-mainz.de
Johannes Gutenberg-Universität
Mainz, Germany

Target Author

[Joëlle Proust](#)

joelle.proust@ehess.fr
Ecole Normale Supérieure
Paris, France

Editors

[Thomas Metzinger](#)

metzinger@uni-mainz.de
Johannes Gutenberg-Universität
Mainz, Germany

[Jennifer M. Windt](#)

jennifer.windt@monash.edu
Monash University
Melbourne, Australia

1 The expressive mode of feelings

First, I would like to repeat, in short, the main claims of the target paper that will serve as a basis for my subsequent comments and extensions in the following sections. Joëlle Proust’s article is concerned with the functional and informational characterisation of feelings. She argues that the concept of “feeling” consists of the following components:

1. Reactive (associated with appraisal)
2. Subjective experience

3. With distinctive embodied phenomenal quality (somatic markers have the function of expressing intensity and valence of feelings, [Proust this collection](#), p. 8)
4. Possessing a formal object (not always, e.g., feeling depressed; absence of a formal object is typical of moods, footnote 5)

The formal object of feelings is argued to be affordance-sensing, a “non-conceptual and entirely subjective appraisal of the environment by the

agent” (Proust [this collection](#), p. 7) or a “*subjective, embodied relation* to an opportunity in an input from the environment” (p. 8). Assuming the non-referential indexicality of feelings, or that feelings signal a *relational* affordance (p. 7) that depends on the representational system (p. 8), Proust argues that feelings can misrepresent only if they are re-described in propositional terms. She argues that feelings are *transparent*, because of the experienced connection between their somatic markers and affordances, as well as because of the direct influence of their valence and intensity on an agent’s motivation (p. 12). Though subjects feel directly, in order to report their feelings they have to “translate one mode of representation into another, with no guarantee that this translation will not enrich or modify FS intentional content” (p. 12). Subjects might reinterpret and mis-describe their feelings, but they cannot thereby change the nature of those feelings (feelings being *cognitively impenetrable*; p. 19).

Feelings are argued to be a plausible candidate for a natural kind on the basis of the comparison between feelings and emotions—which she considers not to constitute a natural kind (Proust [this collection](#), p. 3). Two kinds of subjective appraisal might be part of an emotion: *primary feelings* on the one hand and *appraisals cum conative dispositions* on the other. While the first kind corresponds to an earlier time in our evolutionary development, is independent of concepts, induces specific responses, and possesses distinct somatic markers, the second kind is not and might be a blend of different instances of the first kind. Apart from primary affective feelings, somatic, agential, and metacognitive feelings are argued to form a natural kind.

The function of feelings is to non-conceptually evaluate and signal the result of a comparison process between prediction and outcome through embodied experience (Proust [this collection](#), p. 4). Due to their non-conceptual monitoring nature, feelings do not convey, but merely approximate a causal relation between internal states, external states, and actions (p. 11). There are three kinds of functional relations between feelings and actions (pp. 4–5):

1. Determination of a kind of action in response: approach vs. avoidance
2. Specific orientation in time: predictive vs. retrospective
3. Level of arousal: elevation in energy vs. soothing effect

Feelings are argued to be the result of a *comparator* or control mechanism that is *automatic* and *encapsulated*. The latter requirements are imposed in order to explain the independence of feelings of beliefs and intentions (p. 5) such that, e.g., one could still feel the adrenalin rush even though the hypothesized venomous snake turned out to be a twig.

Metacognitive feelings (M-feelings) are held to express informational, instead of environmental affordances, arise in mental acts, and trigger similar actions of approach or avoidance. M-feelings involve appraisal of the properties of the informational processes underlying contents of thought, but not those content themselves. Against Schachter & Singer’s (1962) two-factor theory of emotions (interpreted as feelings possessing intrinsic arousal but extrinsic valence), Proust argues that feelings have *intrinsic intensity and valence*. Cues on which those feelings are based can be conveyed verbally though, and thus, the heuristics (implicitly or explicitly) might change in the long run. The main claim is thus that *context-dependency is not concept-dependency* (Proust [this collection](#), p. 17). Experience of tokens of feelings differs with respect to the kind of affordance they express (several affordances might be linked to the same type of feeling) and actions they trigger.

An especially interesting claim for me is that affective feeling in general, and metacognitive feelings in particular, have a meta-monitoring function of signalling “the rate of reduction of discrepancy toward a confidence threshold” (Proust [this collection](#), p. 21). If the rate of discrepancy reduction is above expected, the valence of a feeling is experienced as more positive, and, if below expected, as more negative. “Cold” feelings without valence are those for which the expectation has been correct. This claim is interesting for two reasons. On the one hand, to the reader familiar with the self-decep-

tion literature the key-concept “confidence threshold” will stand out. It plays an important role in accounts of self-deception that regard it as a kind of hypothesis testing (one prominent proponent of this view is [Mele 2012](#)). In short, according to this type of account, gathering of evidence in favour of a certain hypothesis is pursued up to a certain point: up until the amount of evidence has reached a confidence threshold that is enough to push an acceptance or rejection of the hypothesis (for more see [Pliushch & Metzinger 2015](#)). On the other hand, “prediction error”, or difference between prediction and sensory input, is the key-term in the model of mental representation that has lately gained a large amount of acceptance—predictive coding (for a short introduction to the free-energy principle of which predictive coding is a particular implementation see [Friston 2009](#); see also [Clark, Hohwy, Seth this collection](#)). Predictive coding provides a unifying explanation for perception, cognition, and action as a result of hierarchical Bayesian inference: at different levels, predictions are compared to propagated precision-weighted prediction error that, under different conditions, leads either to changes in the model of causes of sensory input or to action directed at testing the current model ([Clark 2013](#)).

The idea that feelings signal the rate of reduction of prediction error might be worth elaborating in the predictive coding framework, particularly given the recent study by [Furl et al. \(2010\)](#) who argue that facial expressions are represented as anticipated *trajectories* of the change of those expressions: pictures of neutral and fearful faces were morphed to different degrees such that participants got to see trajectories from a neutral to a fearful face and vice versa. After seeing such a sequences of pictures, participants had to rate another picture for fearfulness. The results indicated that predictable sequences in which the degree of being morphed rose or fell monotonously, thus forming a trajectory, biased perception ([Furl et al. 2010](#), p. 696). Combining Proust’s idea with the results of Furl et al.’s study: feelings might also be represented as anticipated trajectories of change, particularly given the possibly bi-direc-

tional causal influence between feelings and facial expressions (see section 2.2).

2 Critique: Affect and implicit heuristics in feelings

2.1 Use of the term “affect”

The aim of this section is threefold: 1) show difficulties in disentangling feelings from emotions; 2) attempt to give a more precise characterisation of the formal object of feelings, along with some examples; 3) criticize the use of the term “direct” and offer another definition that is consistent with predictive coding. The first problematic point that I see is Proust’s use of the term “affective”, which is ambiguous. She employs at least two different definitions of “affective”:

1. Feelings that possess valence (p. 20). Yet *all* kinds of feelings, according to Proust, possess affect *and* valence¹ (p. 1). Given her distinction between “hot” (emotional) feelings and those that have valence² (p. 21), emotional feelings might differ from mere feelings with valence due to the differently-experienced valence, maybe if emotional valence were a richer experience. Thus, the question is about the minimal requirements on valence and intensity in feelings.
2. Feelings that express emotions.
3. *Difference between feelings and emotions*: if agentive and metacognitive feelings can be affective, then the categorization of feelings into bodily, agentive, metacognitive, *and* affective (p. 5) might be better restricted to the first three, with the fourth being a dimension along which they vary. If affective in this categorization means emotional (p. 2), then there is an ambiguity of terms—affect-

1 The following quotations might help to elucidate the matter: “[f]eelings typically express affect and valence in sensation (25-26), all the feelings vary in affect in roughly the same way, because they all include valence in their informational structure” (p. 20).

2 In Proust’s words, the difference between “hot” feelings and feelings with valence, on the example of M-feelings, is that “although all M-feelings do not often have a definite ‘hot’ quality comparable with fear and love, they always have a valence, according to whether they predict the agent’s progress toward or away from her cognitive goal” (p. 21).

ive = having valence and affective = being part of an emotional experience—because the latter seems to be more complex.

4. *Difference between formal objects of feelings and emotions*: if “feelings are affective ingredients in emotional awareness” (p. 3), then there is a circularity in understanding affectivity here: feelings are affective in virtue of being part of an emotion, while at the same time they themselves are the affective component in the emotion of which they are part. The first part of this claim can be followed from that defended by Proust, namely that feelings that do not express emotions are not affective (p. 2). The second part of the claim follows from Proust’s claim that feelings are affective ingredients of emotions (p. 3). As elaborated in the previous section, emotions are said by Proust to contain one of two kinds of subjective appraisals: feelings or appraisals cum conative dispositions. Further, if feelings are components of emotions, but both can have a formal object, then those objects might diverge. The consequence is that an emotion and a feeling that is part of it might be directed at different objects. Thus, Proust on the one hand distinguishes feelings from emotions and yet on the other hand claims that not only emotional feelings, but also agentive and metacognitive feelings might be “feeling toward” experiences (p. 3, pp. 20–21). The latter claim that both feelings and emotions are directed at intentional objects has been used as an argument to identify both (see [de Sousa 2014](#) section 2 for a discussion of this question). Given Proust’s claim that there are somatic, affective, agentive, and metacognitive feelings, and given the claim that at least in metacognitive feelings the formal object is not the cognitive disposition itself but the rate of change of its execution above or below discrepancy, an interesting question focuses on the formal object of emotional feelings.³ For example, can it be

that while the formal object of the emotion of fear is some dangerous object, the object of a feeling is a rate of change in the assessment of the situation before and after the change of the formal object of an emotion? This might explain why, e.g., the first bite of a bar of chocolate makes one happier than the following bites.

5. *Bodily phenomenology of feelings as their formal object*: [Proust](#) argues that while somatic feelings are about bodily sensations (or, more consistently, about the rate of their change), in affective (emotional) and possibly metacognitive feelings “the bodily phenomenology tends to recede to the fringe of consciousness” ([this collection](#), p. 2). The example that Proust gives with respect to metacognitive feelings is that feelings of remembering are correlated with but not about facial muscle activity (p. 3). Proust acknowledges that there might be mixed cases (experience of bodily feeling + intentional content, pp. 2–3), but I want to argue that in some emotional feelings bodily phenomenology is, to borrow a metaphor, in the *foreground*. There might be emotional feelings whose objects are bodily sensations, e.g., the anxiety that arises during a panic attack: when I concentrate on my accelerated heartbeat, then if I come to associate the heartbeat with some threatening aspects of a situation, such an experience might lead to anxiety, and thus the initial anxiety leads to even more anxiety, leading to a vicious cycle of panic (for a discussion of heartbeat perception in panic disorder see [Ehlers & Breuer 1996](#)). This might be a case of an emotion whose formal object is the rate of change of bodily sensations, or maybe a meta-feeling (for a discussion of meta-emotions see [Mendonça 2013](#)).

In the given panic example it might have seemed as if I had embraced the analogy between feelings and perception that Proust

positive or negative (*ibid.*, p. 7). Positivity and negativity are dimensions along which valence changes, and valence has been characterised as the rate of change of discrepancy towards the (cognitive) goal. For more on why the latter characterisation is interesting see section 3.

³ To be more precise, the question is about the functional description of the formal object of feelings. [Proust](#) ([this collection](#)) says that “[f]eelings express [...] affordance as their focus (for formal object), along with its graded valence, ranging from very unpleasant to very pleasant, and with its intensity gradient, which ranges from small to large” (p. 8). Affordance is defined as “perceived utility”, and can be

denies, so I will explain why it may be more beneficial to use the term “direct” in another sense to that used by Proust. Proust makes a sharp distinction between feeling and perceiving: “[w]hile percepts allow recognition and identification of external objects and properties, feelings express specific affordances in a perceived, imagined, or remembered situation” (this collection, p. 10). Non-conceptual parts of perceptions are said to relate to “objective, external contrastive cues” (Proust this collection, p. 10), while in feelings they relate to evaluative states. Perception is said to involve “direct sensory access to the world” (p. 10), while the access of feelings to the world and the body is claimed to be indirect. Proust’s evidence for a disanalogy between feeling and perception is based on the neuroscientific research of Barrett & Bar, who say that absence of “internal affective context” impairs the categorization of objects (2009, p. 12).⁴ Their evidence for this hypothesis is based on reviewing the anatomic connections involved in affective processing and that of object perception. One critique of this might be that the time of activation of certain regions responsible for emotional processing and perception might justify the claim that emotional processing comes before perception, but not how direct such processing is. Moreover, in light of predictive coding, perception, emotion, and cognition might all be indirect (Hohwy 2014; for more technical elaboration Friston et al. 2014). In other words, predictive coding provides the term “direct” with a meaning other than that used by Proust. In predictive coding directness is an absence of the evidentiary boundary, where the evidentiary boundary is the inferential isolation between the model of the world and the hidden causes of sensory input (Hohwy 2014). This means that causes beyond the boundary have to be inferred on the basis of independent evidence (ibid., p. 6), or, in Hohwy’s words, “[t]he brain doing the inference is secluded at least in the sense that certain

kinds of doubt about the occurrence of the evidence are unanswerable without further, independent evidence” (p. 7). Relating this observation to Proust, on the premise of accepting predictive coding, there might not be a sharp distinction between feeling and perceiving such as Proust postulates, or at least not in the form presented in the target article. If interoception as perception involves inferences about circumstances *beyond* the (same) evidentiary boundary, as suggested by Hohwy (2014), then feeling and perceiving would both be indirect (to the same degree).⁵ If interoception does not go beyond the evidentiary boundary, feelings might be direct, even if perception is not.

2.2 Concept-based feelings?

In this part of the review I will point out the dangers of interpreting the relation between feelings and concepts too simplistically and argue that it is possible that at least some kinds of feelings are influenced by concepts, even if they themselves are non-conceptual. Proust argues that for metacognitive feelings to arise an important affordance, as well as an implicit heuristic, has to be present (this collection, p. 18). This heuristic is based on cues about the dynamics of information processing, but not its contents (p. 15). The dichotomies that Proust uses in the description—implicit–explicit, unconscious–conscious, evolutionarily-old–evolutionarily-new, associative–rule-based (pp. 3–4, p. 17)—have often been mapped onto two different kinds of processes in dual processing theory (e.g., Frankish & Evans 2009). Dual processing theory states that there are two kinds of processing that possess the dichotomous characteristics mentioned above. A minimal description provided by Evans (2009) for type 1 is “fast, automatic, high processing capacity, low effort”, and for type 2 “slow, controlled, limited capacity, high effort” (p. 33). Along these lines, “implicit”, “unconscious”, “evolutionarily old”, “associative” have been also used as descriptors for type 1 and “explicit”, “conscious”, “evolutionar-

⁴ Barrett & Bar (2009) define affect as an influence on bodily states that is either unconscious or, if conscious, experienced as pleasurable or unpleasurable to varying degrees (pp. 1327–1328). Barrett & Bar’s (2009) basic claim is that the orbitofrontal cortex (OFC) integrates into a unified multimodal representation sensory information from both world and body in a dynamic way.

⁵ One could also ask whether the *same* evidentiary boundaries would be involved in feeling and perceiving, since there could be many of them (Friston 2013).

ily new”, “rule-based” as descriptors for type 2. A belief bias (accepting more believable than unbelievable conclusions) might serve as an example for type 1 (*ibid.*, p. 41), and the conscious correction thereof for type 2. The worry I have is adding to those dichotomies another one: non-conceptual (meaning in this case non-propositional; Proust *this collection*, p. 7)–conceptual (propositional, belief-like). Proust holds that “cues (associative heuristics) dictate how an affordance is detected, assessed and exploited in a context, but these cues are not consciously available, and hence do not depend on a naïve theory of the task” (p. 17). This inference is not valid in the given form. I agree with Proust that “[a] cue-based, non-analytic heuristic is not inferential in the interpretive, first-person sense” (p. 17), but I hold that there is at least one step to consider in between non-conceptual⁶ affordances and consciously evaluated affordances. And this is automatic *concept-based* activation (the existence of automatic appraisal is acknowledged by Proust; footnote 7).

Evans (2009) distinguishes between different kinds of dual processing theories, among which are the sequential (first automatic processing, then controlled) and the parallel theory. Proust seems to embrace a sequential kind of dual processing theory, given the functional role she ascribes to metacognitive feelings (evaluation of mental actions before and after their execution; Proust 2013). Yet how far implicit heuristics are independent of concepts is in question. Proust (*this collection*) denies that “a concept-based interpretation will affect the experienced feeling itself” (p. 17). As mentioned in section 1, she also denies that feelings have a conceptual format. Thus, she seems to deny both that concepts play a causal role in the emergence of feelings and that feelings themselves possess a conceptual format. I will briefly demonstrate that the term “implicit heuristic” does not preclude automatic concept activation, if it implies the activation of knowledge or goal representations. Thompson (2009) argues that

heuristic processes are contaminated by background knowledge, as well as by beliefs and expectations (p. 172, p. 174). Frankish (2009) notes that “the concepts of belief and desire correspond to the psychologist’s concepts of knowledge (or memory) and goal structure” (p. 91). Hence, activation of knowledge that may provide the context for feelings could also be conceptual. Goal representations might also be activated in the course of context creation, provided that unconscious goal pursuit is flexible and context-sensitive (Aarts & Custers 2012). Further, unconsciously activated goals not only depend on context, but also *create* context by influencing the accessibility of knowledge, evaluations, and emotions (Fishbach & Ferguson 2007, p. 496). It follows that if goal representations are activated, then they might lead to the activation of conceptual knowledge. Another interesting point is that if there is a continuous interplay between goal representations and affordances (opportunities in the environment; Huang & Bargh 2014, p. 125) and if goal representations can change the experience of the world (*ibid.*, p. 124), then goal representations might change sensing of affordances and, hence, the feelings associated with it. Further, there has been a proposal to distinguish between associative and rule-based processes by the kind of architecture they operate upon: namely connectionist vs. classical computational (for a short discussion see Samuels 2009, pp. 141–142). Thus, implicit heuristics might be understood as certain connected representations in a network being activated by some cues, where the question is about the representational format of such knowledge, or a more precise description of the relational nature of the feeling affordance. Last, a general note about the similarity between feelings and other kinds of representations: if Bliss-Moreau & Williams (2014) are correct in defending the claim that all kinds of representations possess an affective component (valence + arousal in their definition), then affect is something that expressive and conceptual representations share.

Of course, Proust’s claim that in the case of feelings those cues relate to the dynamics, but not to the contents of processes, indicates a

⁶ Among those who agree with Proust that the content of epistemic feelings is non-conceptual and non-metarepresentational are, for example, Michaelian & Arango-Muñoz (2014). But the *content* of a metacognitive feeling being non-conceptual does not preclude that concepts play a causal role in its emergence.

more specific understanding of the kind of implicit heuristic in question. My point, though, is that if humans can “enrich their noetic feelings through concepts, and thereby revise their reliance on fluency where it is not justified” (Proust 2013, p. 144), then in humans implicit heuristics may also be influenced by concepts (in an automatic way) and in such a way influence feelings. Needless to say, the independent existence of such a schema (be it cognitive or emotional) is hard to prove (Eysenck & Keane 2010, p. 597). According to Koriat & Levy-Sadot (1999), as cited by Proust (this collection, p. 15), metacognitive feelings arise as a result of nonanalytic inferential processes (described as the implicit or unconscious application of heuristics), in distinction to the direct memory trace hypothesis, according to which feelings have direct access to memory traces (Koriat & Levy-Sadot 1999, p. 487). Koriat & Levy-Sadot (1999) argue that the presence of dissociations between knowing and the feeling of knowing speaks against the second hypothesis. Even if heuristics in feelings are non-conceptual, the fact that through feelings emotion gets its valence necessitates that we consider how concepts and memory traces influence feelings, given that they play a role in emotions. Lane et al. (forthcoming), for example, argue that psychotherapeutic change is made possible by updating prior emotional experiences, for which memory traces of those experiences have to be reactivated and reconsolidated. Thus, even if feelings are non-propositional (Proust this collection, p. 20), activation of concepts and their expression in propositional terms are to be distinguished. The point is not that metacognitive feelings themselves cannot have indexical formats,⁷ or that an agent could not possess expressive and conceptual representations at the

same time, but that in humans the generation of (at least) metacognitive and emotional feelings might be preceded by an automatic concept activation that influences them. If this is the case, then one could ask again whether feelings are transparent (see section 1).

Further, instead of describing cognitive processes as serial, their *dynamic* (continuous) nature might be more worthy of emphasis. In the target article, Proust mentions that “[i]ncreased activity in the smile muscle, the zygomaticus major, produces feelings with a positive valence” (this collection, p. 15). This suggests that facial expression influences emotions. She also argues for the transparency (impenetrable nature) of feelings and the against two-factor theory, thus against the possibility that appraisal influences the valence of feelings (see section 1). I want to offer for clarification purposes a short review of the recent literature on which factors are supposed to influence feelings and factors feelings influence themselves. Rogers et al. (2014) emphasize the dynamic nature of emotions insofar as they depend on the social appraisal of a situation. Brosch (2013) also emphasizes the dynamic nature of appraisal that plays a causal role in eliciting emotions. The definition of appraisal that Brosch (2013) accepts also encompasses low-level appraisal based on learned schemata (p. 370). Brosch (2013) argues that first an initial low-level appraisal affects the physiology (1), action tendency (2), expression (3), and feeling (4) of an emotional experience, and then those changes in turn affect an on-going (low- and high-level) appraisal, establishing an appraisal loop. Here, the direction of influence is still in question, e.g., whether feelings influence expressions or the other way around. Laird & Lacasse (2014) defend the James-Lange theory of emotion, namely that facial expressions (e.g., BOTOX patients being less responsive to mild positive emotional stimuli; for the reference see *ibid.*, p. 29), expressive behaviour (e.g., romantic attraction as a result of shared, mutual gaze; *ibid.*, p. 29), and visceral responses that are interpreted according to situational cues (e.g., misattribution of emotion) are *causes* of emotions (for a critique of their evidence see Reisenzein &

⁷ A better understanding of the indexical mode of feelings might be provided by the following quotation: “Feelings can be seen as pre-specified states of a comparator, which predict ultimate success or failure in the actions that they monitor. Given that the information they carry is immediately used in controlling and monitoring current effort, it is misleading to present them as ‘reporting’ the epistemic properties of a mental state or referring to it (even *de re*). They are, rather, signals in a control mechanism, which work somewhat as traffic lights do: allowing traffic, stopping it, rechanneling it; no report or reference need be involved” (Proust 2013, p. 76). In another place Proust (2013) notes that feelings “do not properly ‘refer’, because they do not engage propositional thinking” (p. 77).

Stephan 2014). As such, they may influence emotional feelings too, which Proust acknowledges by pointing out the causal connection between measures in facial muscles and affective character of feelings (this collection, p. 25). Yet the direction of influence may also go the other way around (from feelings to facial expressions). Thus, the nature of feelings may also be dynamic, as are the nature of the underlying cognitive processes. Interestingly, Thagard & Schröder (2014) argue for a neurocomputational theory of emotions as semantic pointers (term introduced by Chris Eliasmith). They argue that physiological, appraisal, social, and psychological components of emotions can be integrated into one unified account: emotion tokens can possess both shallow and deep meanings. The compressed (shallow) form of emotions is reportable, while at the same time pointing to the uncompressed deep form that binds together situational, physiological, and appraisal components.

In the preceding paragraph I considered literature supporting the claim that feelings are embedded in *continuous* cognitive processes. The purpose of this was to show that how appraisal might influence feelings in some form is complex and might even be circular. In this paragraph I offer some additional evidence against a discontinuous interpretation of the connection between feelings and propositional descriptions thereof. The existence of *affective blindsight* (ability to discern emotional stimuli despite inability to consciously perceive them; Eysenck & Keane 2010, p. 581) would stand in line with the assumption that emotional and cognitive processing is based on different kinds of information. This is because affective blindsight demonstrates the dissociation between two different kinds of processing and, thus, a dissociation between the information needed for the one kind and for the other. Further, Scott et al.'s (2014) experiment demonstrating *blind insight* (accurate metacognitive accuracy in the absence of discriminative accuracy) on the one hand supports Proust's hypothesis that metacognition and first-order cognition are not based on the same kind of information, yet on the other it speaks against a serial interpretation

according to which feelings arise out of automatic processes and are then re-described in propositional terms and used in first-person inferential reasoning. Liu & Wang (2014), for example, argue that motivational intensity influences the effect of positive affect on cognitive control: low-approach motivated positive affect enhances cognitive flexibility and distractibility, while high-approach motivated positive affect (associated with goal pursuit) enhances cognitive stability. Thus, the role of feelings might be broader than just the indicators that may or may not be used in conscious reasoning.

3 Proposals: Tension in self-deception is a kind of metacognitive feeling

Proust (this collection, as well as 2013) argues that mental actions are preceded and followed by metacognitive feelings indicating the appropriateness of the cognitive process in question. I want to argue that tension in self-deception fits the characterisation of a metacognitive feeling. Tension is described as a feeling of uneasiness and distress, and as such I think that it is precisely this tension that is said to indicate to the self-deceiver that her belief-forming process is faulty.

Self-deception (SD) is a motivated (1) kind of typically subpersonal hypothesis-testing (2) that results in an evidence-incompatible mental representation of reality (3) which fulfils a belief-like role (4) (Pliushch & Metzinger 2015). Self-deception is usually discussed in the context of biased belief-forming processes and it is argued that phenomenological tension arises as a result of the execution of such processes (e.g., Lynch 2012). Thus, the same function has been ascribed to tension in self-deception as the one ascribed by Proust to metacognitive feelings, namely a comparison of the cognitive process to certain criteria. In self-deception, rationality criteria are typically emphasised.

I want to argue that metacognitive feelings apply to self-deception, insofar as they might also monitor *unconscious* cognitive processes and arise not only before or after a cognitive process, but also *during* it. In case of self-deception these cognitive processes are belief-forming

processes. Proust (this collection, 2013) considers conscious mental actions: her argument is that unconscious comparison processes that give rise to metacognitive feelings precede and follow conscious mental actions. She argues that the “attentional-supervisory system” emerges from “distributed metacognitive abilities” (Proust 2013, p. 263). Ignorance of epistemic norms such as relevance, coherence, fluency, and informativeness lead to (pathological) errors in belief acquisition (Proust 2013, pp. 260–261). My argument in favour of the extension of metacognitive feelings to monitor unconscious cognitive processes is of a phenomenological nature. I agree with Proust (this collection) that the term “inference” has been used loosely in the literature and does not always indicate a first-person inference (p. 21). Yet the more basic problem might be that there is no sequential first-person inference as such in the first place. If the shift between mind wandering (task-unrelated cognitive activity) and task-directed cognitive activity goes unnoticed (Metzinger 2013), then there might be other shifts that we do not notice, e.g., the shift from unconscious to conscious cognitive processes, or some changes in the given process. Thus, the phenomenology of a cognitive process might be more complicated than a unified sequence with a starting point and an end. Further, given, for example, mood-state dependent cognition (Eysenck & Keane 2010, pp. 584), I doubt the plausibility of the assumption that only in breaks between conscious cognitive processes do subjects experience affective feelings.

In the previous paragraphs I argued that the functional role of metacognitive feelings fits that of tension in self-deception, and that metacognitive feelings arise not only before and after mental actions, but also before, after, and during unconscious (possibly self-deceptive) cognitive processes. In this paragraph I want to link Proust’s idea that feelings possess valence only if the rate of change of progress is unexpected to predictive coding, in order to provide a functional description of metacognitive feelings. Proust (this collection) argues that the affective quality of feelings arises only if the cognitive process violates expectations: if it progresses

quicker towards the goal, positive feelings arise, if slower, negative feelings arise⁸ (p. 21). Given that the terms “expectation” and “prediction error” have gained popularity in virtue of being key terms in predictive coding, which is a modelling strategy explaining perception, cognition, and action (Clark 2013), I will shortly discuss Proust’s claim about affect in metacognitive feelings in the context of predictive coding. According to predictive coding, prediction errors (deviation between expectation and outcome) are precision-weighted. Precision is the property of prediction errors (errors between the top-down prediction and the bottom-up signal one receives) that can be described as the weight of a prediction error that plays the role of selection: the more precise the prediction error, the more it will change the hypothesis about causes of input. Switching between perception and action depends on the precision of prediction errors: precise prediction errors change hypotheses, while imprecise ones lead to action (Brown et al. 2013). Precision⁹ is also argued to play a dual biasing role: biasing perception toward goal states and enhancing confidence in action choices (Friston et al. 2013). Low precision of prediction errors has been argued to cause anxiety (Mathys et al. 2011, p. 17).¹⁰ I argue that Proust’s proposal that violations of expectations of “a given rate of reduction of the discrepancies toward her [agent’s] cognitive goal” (this collection, p. 26) produce affective feelings might be described in predictive coding terms as violations of *transition probabilities* of reaching the goal state:¹¹ if a state conducive to the goal state or a goal state itself has been reached, despite a low probability of changing

8 Note the analogy to the “dark room problem” in predictive coding: if an agent wants to minimize surprise or prediction error, then she should stay in a dark room, given that there will be no surprise in it (e.g., Clark 2013). If there were no prediction error, this would cause uncertainty (e.g., Friston et al. 2012). Proust’s argument is similar: if there were no violations of expectations, then metacognitive feelings would not have any valence, because they only have valence if the rate of change is quicker or slower than expected.

9 Attention is precision optimization according to predictive coding (Hohwy 2013).

10 Mathys et al. (2011) are also interesting for the given topic insofar as Proust argues that the heuristics upon which metacognitive feelings are based might be changed via associative learning; Mathys et al. (2011) provide a predictive coding model of reinforcement learning.

11 For a predictive-coding model of a goal-directed action see Friston et al. (2013).

into that state from the current state, then positive affective feelings might arise.¹²

The first step in the categorisation of tension as a metacognitive feeling has been an extension of the application of metacognitive feelings to unconscious belief-forming processes. The second is to clarify the representational content of tension. To do the latter, it might be beneficial to consider which other kinds of metacognitive feelings arise out of belief-forming processes. Those are intuitivity, counter-intuitivity, and anxiety, if one classifies them according to the phenomenology and not according to the norm that they control. Intuitivity indicates the appropriateness of a given belief-forming process.¹³ The reason for the ascription of the given functional role to intuitivity is that intuitivity signals 1) a good fit with respect to the network of our explicit background beliefs and 2) a good fit with respect to our conscious and unconscious model of reality (Metzinger & Windt 2014). An appropriate belief-forming process provides a good fit with respect to 1) and highly likely also with respect to 2). I further argue that counter-intuitivity represents that a certain cognitive process violates the chosen criterion of appropriateness, but is neutral with respect to the system's goal representations, while tension or anxiety represents that the cognitive process violates at least some important goal representations. The reason for this distinction is to account for the effect of motivation on belief-forming processes.

Thus, if feelings accompany our belief-forming processes, then readers might have experienced some while reading this commentary: hence the title. To conclude, I think that Proust has offered interesting ideas on the nature of feelings that will greatly contribute to the clari-

fication of the matter: the indexical (affordance-sensing and non-conceptual) format of feelings, their transparency, the taxonomy of feelings into sensory, emotional, agentive, and epistemic, the predictive and retrospective function of feelings signalling the appropriateness of the cognitive process they monitor, and the degree of change of expectation as the origin of valence of feelings. In this review I have tried to extend Proust's account. To do this, I attempted to provide some conceptual clarifications on the distinction between feelings and emotions, the formal object of feelings, and the conceptual influences to which they might be subject. Last, I argued that tension in self-deception is a kind of metacognitive feeling.

Acknowledgements

I am very grateful to the Barbara-Wengeler foundation for generously supporting my PhD project and all the reviewers of this commentary for their helpful improvements.

¹² Emotional valence has been also argued to be modelled as the rate of change of free energy: Instead of estimating volatility or “slow and continuous changes in states of the world” the rate of change of free energy is argued to take that role of estimating (known) uncertainty (Joffily & Coricelli 2013, p. 1). Here Joffily & Coricelli (2013) accept Yu & Dayan's (2005) distinction between expected and unexpected uncertainty: Expected uncertainty is the one about known unreliability of predicting relationships *within* a context and unexpected uncertainty is the one about the appropriateness of the context itself such that when unexpected uncertainty is high, it is a signal that a *context switch* should be made.

¹³ For an elaboration on the phenomenal signature of knowing in intuitions of certainty, see Metzinger & Windt (2014).

References

- Aarts, H. & Custers, R. (2012). Unconscious goal pursuit: Nonconscious goal regulation and motivation. In R. M. Ryan (Ed.) *The Oxford handbook of human motivation* (pp. 232-247). Oxford, UK: Oxford University Press.
- Barrett, L. F. & Bar, M. (2009). See it with feeling: Affective predictions during object perception. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364, 1325-1334.
[10.1098/rstb.2008.0312](https://doi.org/10.1098/rstb.2008.0312)
- Bliss-Moreau, E. & Williams, L. A. (2014). Tag, you're it: Affect tagging promotes goal formation and selection. *Behavioral and Brain Sciences*, 37 (2), 138-139.
[10.1017/S0140525X13001969](https://doi.org/10.1017/S0140525X13001969)
- Brosch, T. (2013). Comment: On the role of appraisal processes in the construction of emotion. *Emotion Review*, 5 (4), 369-373. [10.1177/1754073913489752](https://doi.org/10.1177/1754073913489752)
- Brown, H., Adams, R. A., Parees, I., Edwards, M. & Friston, K. (2013). Active inference, sensory attenuation and illusions. *Cognitive Processing*, 14 (4), 411-427.
[10.1007/s10339-013-0571-3](https://doi.org/10.1007/s10339-013-0571-3)
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36 (3), 181-204.
[10.1017/S0140525X12000477](https://doi.org/10.1017/S0140525X12000477)
- Clark A. (2015). Embodied prediction. In T. Metzinger and J. M. Windt (Eds.) *Open MIND*. Frankfurt a. M., GER: MIND Group.
- de Sousa, R. (2014). Emotion. In E. N. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy (Spring 2014 Edition)*.
<http://plato.stanford.edu/archives/spr2014/entries/emotion/>
- Ehlers, A. & Breuer, P. (1996). How good are patients with panic disorder at perceiving their heartbeats? *Biological Psychology*, 42 (1-2), 165-182.
[10.1016/0301-0511\(95\)05153-8](https://doi.org/10.1016/0301-0511(95)05153-8)
- Evans, J. St. B. T. (2009). How many dual-process theories do we need? One, two, or many? In K. Frankish and J. St. B. T. Evans (Eds.) *In two minds: Dual processes and beyond* (pp. 33-54). Oxford, UK: Oxford University Press.
- Eysenck, M. & Keane, M. T. (2010). *Cognitive psychology: A student's handbook*. Hove, UK: Psychology Press.
- Fishbach, A. & Ferguson, M. (2007). The goal construct in social psychology. In A. W. Kruglanski and E. T. Higgins (Eds.) *Social psychology: Handbook of basic principles* (pp. 490-515). New York, NY: Guilford Press.
- Frankish, K. (2009). Systems and levels: Dual-system theories and the personal-subpersonal distinction. In K. Frankish and J. St. B. T. Evans (Eds.) *In two minds: Dual processes and beyond* (pp. 89-107). Oxford, UK: Oxford University Press.
- Frankish, K. & Evans, J. St. B. T. (2009). The duality of mind: An historical perspective. In K. Frankish and J. St. B. T. Evans (Eds.) *In two minds: Dual processes and beyond* (pp. 1-29). Oxford, UK: Oxford University Press.
- Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences*, 13 (7), 293-301. [10.1016/j.tics.2009.04.005](https://doi.org/10.1016/j.tics.2009.04.005)
- (2013). Life as we know it. *Journal of The Royal Society Interface*, 10 (86), 20130475.
[10.1098/rsif.2013.0475](https://doi.org/10.1098/rsif.2013.0475)
- Friston, K., Adams, R. A., Perrinet, L. & Breakspear, M. (2012). Perceptions as hypotheses: Saccades as experiments. *Frontiers in Psychology*, 3 (151).
[10.3389/fpsyg.2012.00151](https://doi.org/10.3389/fpsyg.2012.00151)
- Friston, K., Schwartenbeck, P., Fitz-Gerald, T., Moutoussis, M., Behrens, T. & Dolan, R. J. (2013). The anatomy of choice: Active inference and agency. *Frontiers in Human Neuroscience*, 7 (598).
[10.3389/fnhum.2013.00598](https://doi.org/10.3389/fnhum.2013.00598)
- Friston, K., Sengupta, B. & Auletta, G. (2014). Cognitive dynamics: From attractors to active inference. *Proceedings of the IEEE*, 102 (4), 427-445.
[10.1109/JPROC.2014.2306251](https://doi.org/10.1109/JPROC.2014.2306251)
- Furl, N., van Rijsbergen, N.J., Kiebel, S. J., Friston, K. J., Treves, A. & Dolan, R. J. (2010). Modulation of perception and brain activity by predictable trajectories of facial expressions. *Cerebral Cortex*, 20 (3), 694-703.
[10.1093/cercor/bhp140](https://doi.org/10.1093/cercor/bhp140)
- Hohwy, J. (2013). *The predictive mind*. Oxford, UK: Oxford University Press.
- (2014). The self-evidencing brain. *Noûs*.
[10.1111/nous.12062](https://doi.org/10.1111/nous.12062)
- Hohwy, J. (2015). The neural organ explains the mind. In T. Metzinger and J. M. Windt (Eds.) *Open MIND*. Frankfurt a. M., GER: MIND Group.
- Huang, J. Y. & Bargh, J. A. (2014). The selfish goal: Autonomously operating motivational structures as the proximate cause of human judgment and behavior. *Behavioral and Brain Sciences*, 37 (2), 121-135.
[10.1017/S0140525X13000290](https://doi.org/10.1017/S0140525X13000290)
- Joffily, M. & Coricelli, G. (2013). Emotional valence and the free-energy principle. *PLoS Computational Biology*, 9 (6), e1003094. [10.1371/journal.pcbi.1003094](https://doi.org/10.1371/journal.pcbi.1003094)

- Koriat, A. & Levy-Sadot, R. (1999). Processes underlying metacognitive judgments: Information-based and experience-based monitoring of one's own knowledge. In S. Chaiken and Y. Trope (Eds.) *Dual-process theories in social psychology* (pp. 483-502). London, UK: The Guilford Press.
- Laird, J. D. & Lacasse, K. (2014). Bodily influences on imotional feelings: Accumulating evidence and extensions of William James's theory of emotion. *Emotion Review*, 6 (1), 27-34. [10.1177/1754073913494899](https://doi.org/10.1177/1754073913494899)
- Lane, R. D., Ryan, L., Nadel, L. & Greenberg, L. (forthcoming). Memory reconsolidation, emotional arousal and the process of change in psychotherapy: New insights from brain science. *Behavioral and Brain Sciences*. [10.1017/S0140525X14000041](https://doi.org/10.1017/S0140525X14000041)
- Liu, Y. & Wang, Z. (2014). Positive affect and cognitive control: Approach-motivation intensity influences the balance between cognitive flexibility and stability. *Psychological Science*, 25 (5), 1116-1123. [10.1177/0956797614525213](https://doi.org/10.1177/0956797614525213)
- Lynch, K. (2012). On the "tension" inherent in self-deception. *Philosophical Psychology*, 25 (3), 433-450. [10.1080/09515089.2011.622364](https://doi.org/10.1080/09515089.2011.622364)
- Mathys, C., Daunizeau, J., Friston, K. J. & Stephan, K. E. (2011). A Bayesian foundation for individual learning under uncertainty. *Frontiers in Human Neuroscience*, 5 (39). [10.3389/fnhum.2011.00039](https://doi.org/10.3389/fnhum.2011.00039)
- Mele, A. R. (2012). When are we self-deceived? *Human-a.Mente - Journal of Philosophical Studies*, 20, 1-15.
- Mendonça, D. (2013). Emotions about emotions. *Emotion Review*, 5 (4), 390-396. [10.1177/1754073913484373](https://doi.org/10.1177/1754073913484373)
- Metzinger, T. (2013). The myth of cognitive agency: Subpersonal thinking as a cyclically recurring loss of mental autonomy. *Frontiers in Psychology*, 4 (931). [10.3389/fpsyg.2013.00931](https://doi.org/10.3389/fpsyg.2013.00931)
- Metzinger, T. & Windt, J. (2014). Die Phänomenale Signatur des Wissens: Experimentelle Philosophie des Geistes mit oder ohne Intuitionen? In T. Grundmann, J. Horvath and J. Kipper (Eds.) *Die experimentelle Philosophie in der Diskussion* (pp. 279-321). Berlin, GER: Suhrkamp.
- Michaelian, K. & Arango-Muñoz, S. (2014). Epistemic feelings, epistemic emotions: Review and introduction to the focus section. *Philosophical Inquiries*, 2 (1), 97-122.
- Pliushch, I. & Metzinger, T. (2015). Self-deception and the dolphin model of cognition. In R. Gennaro (Ed.) *Disturbed consciousness*. Cambridge, MA: MIT Press.
- Proust, J. (2013). *Philosophy of metacognition: Mental agency and self-awareness*. Oxford, UK: Oxford University Press.
- Proust J. (2015). The representational structure of feelings. In T. Metzinger and J. M. Windt (Eds.) *Open MIND*. Frankfurt a. M., GER: MIND Group.
- Reisenzein, R. & Stephan, A. (2014). More on James and the physical basis of emotion. *Emotion Review*, 6 (1), 35-46. [10.1177/1754073913501395](https://doi.org/10.1177/1754073913501395)
- Rogers, K. B., Schröder, T. & Scheve, C. v. (2014). Dissecting the sociality of emotion: A multilevel approach. *Emotion Review*, 6 (2), 124-133. [10.1177/1754073913503383](https://doi.org/10.1177/1754073913503383)
- Samuels, R. (2009). The magical number two, plus or minus: Dual-process theory as a theory of cognitive kinds. In K. Frankish and J. St. B. T. Evans (Eds.) *In two minds* (pp. 129-146). Oxford, UK: Oxford University Press.
- Schachter, S. & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, 69 (5), 379-399. [10.1037/h0046234](https://doi.org/10.1037/h0046234)
- Scott, R., Dienes, Z., Barrett, A. B., Bor, D. & Seth, A. K. (2014). Blind insight: Metacognitive discrimination despite chance task performance. *Psychological Science*, 25 (12), 2199-2208. [10.1177/0956797614553944](https://doi.org/10.1177/0956797614553944)
- Seth A. (2015). The cybernetic bayesian brain. In T. Metzinger and J. M. Windt (Eds.) *Open MIND*. Frankfurt a. M., GER: MIND Group.
- Thagard, P. & Schröder, T. (2014). Emotions as semantic pointers: Constructive neural mechanisms. In L. F. Barrett and J. A. Russell (Eds.) *The psychological construction of emotion* (pp. 144-167). New York, NY: Guilford.
- Thompson, V. A. (2009). Dual-process theories: A metacognitive perspective. In K. Frankish and J. St. B. T. Evans (Eds.) *In two minds: Dual processes and beyond* (pp. 171-195). Oxford, UK: Oxford University Press.
- Yu, A. J. & Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron*, 46 (4), 681-692. [10.1016/j.neuron.2005.04.026](https://doi.org/10.1016/j.neuron.2005.04.026)

Feelings as Evaluative Indicators

A Reply to Iuliia Pliushch

Joëlle Proust

These responses aim at clarifying various aspects and implications of my proposal that feelings are affordance sensings. Affective quality, in the present proposal, extends beyond the domain of primary and secondary emotions to all feelings, because it results from specific features in the dynamics of valence. Feelings do not convey an explicit causal information about the world. Causal relations are, rather, implicitly represented in a felt affordance through the dynamic relations between the associated, embodied cues for location, valence and intensity and type of the affordance. Affordances are neither perceived nor inferred; they are “sensed”, which is an ability distinct from belief, whose informational input is derived from features of a perceived or interpreted situation or cognitive task. The input for an affordance sensing can well be conceptual; it is claimed, however, that even when a task is represented through concepts, the affordance-sensings elicited during the task are nonconceptual and evaluative. The relevant properties in affordance-sensings being dynamic, an interpretation of the view under discussion as being serial is resisted. Finally, Pliushch’s proposal for extending this theory to an interpretation of the feelings involved in self-deception is discussed.

Keywords

Affective feelings | Causal information | Metacognition | Noetic feelings | Self-deception | Serial vs. dynamic processes | Valence

Author

Joëlle Proust

joelle.proust@ehess.fr

Ecole Normale Supérieure
Paris, France

Commentator

Iuliia Pliushch

pliushi@students.uni-mainz.de

Johannes Gutenberg-Universität
Mainz, Germany

Editors

Thomas Metzinger

metzinger@uni-mainz.de

Johannes Gutenberg-Universität
Mainz, Germany

Jennifer M. Windt

jennifer.windt@monash.edu

Monash University
Melbourne, Australia

1 Use of the term “affect”

One of the aims of this article is to try to define feelings according to their functional characteristics, when seen as all-purpose comparators. Iuliia Pliushch claims that my use of “affective feelings” is ambiguous, because they seem to be *defined* either as “feelings that possess valence”, or as “feelings that express emotions”. I am happy to accept the blame for not rephrasing in my own terms the subcategory of “affective feelings” discussed in emotion theory.

A similar discrepancy, however, may seem to be present between two passages of my chapter where I do express my own view:

As will be seen below, some feelings, however, do not express emotions, i.e., are not affective. ([Proust this collection](#), p. 2)

All the feelings vary in affect in roughly the same way, because they all include valence in their informational structure. ([Proust this collection](#) p. 20)

The discrepancy is only apparent, however, and should disappear when the issue of valence in its relation to affect is properly addressed. In emotion theory, the relations between valence and affect, and even the existence of valence,

are highly debated. With rare exceptions,¹ the question is ignored by theorists of somatic, agentive, or noetic feelings.² The proposal summarized in (2), however, posits that affect will result from valence (not the other way round). Section 7 aims to explain why affect depends on the dynamics of valence throughout the domain of feelings. These relations are modulated by the dynamic conditions that prevail in the contrast between expectancy and observation in a given domain. When observation and expectancy coincide with a predicted temporal pattern - with a small stake involved-, the corresponding feelings should not involve affect on top of valence. This is the case for the feelings of agentive success that are generated in routine actions. Hence (1) holds. When you predictably overcome a minor obstacle, you don't feel particularly thrilled. When special dynamic conditions obtain, however, (acceleration or deceleration in the rate of observed change, as compared with the expected rate of change), valence will be intensely felt, in terms of vividly positive or negative experiences. Scoring an ace in a tennis game, especially if it is a rare achievement for this player, elicits in him/her an intensive positive affect. Dynamic variations of this kind also apply to metacognition, where Archimedes' "Eureka" is affect-laden, while the felt ability to respond to a memory question in a laboratory is not.

Hence there may be affect-laden feelings beyond the domain of what is traditionally called "emotional" or "affective feelings". Reciprocally, one might suspect that in the latter domain, too, affect only appears beyond thresholds of positive or negative valence, with colder kinds of feelings occupying the lower end of the continuum.

2 Causal information: Explicit versus implicit

Iuliia Pliushch presents my view on the role of causal relations in feeling representations as fol-

lows: "Due to their non-conceptual monitoring nature, feelings do not convey, but merely approximate a causal relation between internal states and actions" ([this collection](#), p. 2). It may be useful to briefly comment on this summary, in order to clarify the aim of the passage where this question is discussed as follows:

Clearly, FS does not explicitly convey a causal relation between situation, somatic markers and subjective feeling. It carries this causal relation implicitly, however, as a consequence of the control architecture that produces feelings. In an emotional control loop, a perceived affordance causes (not: is represented as causing) its expressive evaluation through its specialized sensory feedback. Emotional awareness expresses this functional relation. ([Proust this collection](#), p. 11)

What is at stake is not the causal relation between internal states and actions, but rather the nature of the causal relation between, on the one hand, the agent's perceptual belief about an external situation ("there is a bear in front of me") and his/her own bodily changes (pounding heart, trembling legs, etc.). According to cognitivists, this causal relation is not only generating a specific emotion, or in my terms, a given feeling, as most theories would accept. It also constitutes in part the intentional content of the experience of fear, or more generally, of an emotional experience. What I object to here is that the *representational* structure of feelings is not *constituted* by a conceptual representation of the causal link between an external fact and observed bodily changes. The causal relations are, rather, implicitly represented in a felt affordance through the dynamic relations between the associated, embodied cues for location, valence and intensity and the type of affordance perceived. Perceiving a bear elicits a bear-affordance (i.e., a feeling of fear of this bear). Even though, from an external viewpoint, one might say that identifying an object as dangerous has caused a disposition to act in the agent, from the viewpoint of the engaged agent, no such judgment needs to be

¹ In particular Carver & Scheier (1990, 2001) for feelings of agentive success or failure, and Stepper & Strack (1993) for noetic feelings.

² For an interesting philosophical discussion of the nature of valence, see Prinz (2010), against Solomon's skeptical stance (2003).

formed because the representation of a given affordance includes the relevant “causal” information in its associative dynamic structure. As suggested by Pliushch, being evaluative, feelings predispose to act adaptively. A disposition to act, then, is associated with an affordance, and with the bodily markers for valence and intensity constituting this affordance.

3 Phenomenology of feelings: Background or foreground?

Should we construe the phenomenology of feelings – the presence of a bodily change – as being in the foreground or in the background of consciousness? The article under review briefly discusses this issue (Pliushch [this collection](#), pp. 2-3): A feeling tends to be more explicitly felt as bodily when making a bodily need salient (feeling tired, feeling a pain in the joints), plausibly because its function is to motivate bodily-directed action. Although in so-called “affective feelings” ¹ the bodily phenomenology tends to recede to the fringe of consciousness, there are cases, as Iuliia Pliushch notes correctly, where it occupies center stage – think of Proust’s report about his chest pain when learning that Madame de Guermantes just died.

It is debatable, however, that in such cases, the formal object of the feeling consists merely in the bodily changes, say, in heartbeat rate. For such states are part of an intensifying negative affordance: the loss of a friend. The notions of “meta-emotion” and “meta-feeling”, which are used by Pliushch to discuss the amplification of a feeling might be captured either in purely dynamic terms, or in a conceptual reconstruction of the situation at hand. This interesting issue, discussed in section 2.2 of Iuliia Pliushch’s comments, has connections with the notion of how concepts and feelings interact, and will be addressed in section 4.

4 Directedness

Iuliia Pliushch objects to my distinction between perceptions and feelings. The claim that “feelings do not have a direct sensory access to the world”, she says, relies on a meaning

of “direct” that is not compatible with the view defended by predictive coding theorists, where “directness is an absence of the evidentiary boundary” (Pliushch [this collection](#), p. 5). Being direct, then, if I understand this sentence correctly, means to lack independent evidence about the world of the kind that perception could bring. Although predictive coding offers a stimulating scheme for understanding mental function, it is open to interpretation and controversy. The functional hypothesis that perceiving and feeling are both indirect will appear highly counter-intuitive to many psychologists and philosophers.

As far as my article is concerned, I have defended the view that feelings are directly related to an opportunity, in the sense that they represent it in an immediate way, a view that has been defended by most affordance theorists. This is compatible with the claim that their informational pathway is derived from perception or memory. What may appear puzzling in my proposal is that an affordance is neither directly *perceived* nor *inferred*. It is directly *sensed*, which requires a different kind of ability. In section 5.1, I have proposed to distinguish associations from inferences, which is relevant to the present discussion. The kind of trigger for feelings are cues elicited in a currently active context, not inferences. These cues are delivered by sensory perception or by memory, but dealt with in a separate subsystem.

5 What are the relations between feelings and conceptual representations?

The comments in section 2.2 of my reviewer’s contribution are presented as an alternative approach to my own view, but I find myself in agreement with most of the claims, in particular with the remarks on p. 6 concerning the relations between feelings and conceptual representations. The main point concerns how one’s own goal, when acting, may influence the production of particular feelings. I discuss this issue at length in sections 5 and 6 of the article under review (Proust [this collection](#)), as well as in a recent publication devoted to action representa-

tions (Proust 2014). My position is captured by two claims. 1) Feelings – affordance sensings – can be, and indeed are usually triggered while performing a task that has been defined in conceptual terms. Cognitive affordances, in particular, are important relational properties that an agent needs to use when attempting to solve highly complex problems, for example when playing chess or looking for a mathematical proof. 2) The feeling episode, however, has an exclusively evaluative, non-conceptual content. I am aware that these two claims may easily be misunderstood. To disentangle the two, think of what agents mean to do: they mean to play chess according to the rules, or to prove a theorem. These goals, indeed, are conceptually represented, and depend on background beliefs and a sensitivity to epistemic norms such as truth and coherence, which presupposes in these particular cases an ability to represent beliefs as beliefs. Feelings of knowing, feelings of being right, and other affordance sensings are generated while the agents are conducting these higher-level forms of reasoning. They are dependent on the mental and neural activity which is thereby elicited. In other words, these feelings do not result from *a consideration* of the concepts involved, but from the dynamic features of the underlying processes. Hence, I would go farther than my reviewer, when she claims that noetic feelings are often elicited when concepts are automatically activated when forming a cognitive goal: they are also elicited when concepts are activated in a controlled way, e.g., in the process of planning what to do.

Should we conclude from this claim that heuristic processes are “contaminated by background knowledge” (Pliushch this collection, p. 6)? No. One should rather conclude that while the goal of a mental action is conceptually defined, the feelings entertained while acting are generated not by the concepts themselves, but by the dynamic characteristics of the processes underlying concept use. It is thus perfectly coherent to conclude that feelings have their own representational format that is not itself “infected” by concepts. A “theory of the task” is not a constituent of an affordance sensing, it is only a precondition for evaluating one’s ability in solving a task.

6 Serial versus dynamic properties of cognitive processes

My reviewer attributes to me a serial view of cognitive processes because I distinguish predictive from retrodictive evaluations of mental actions (Pliushch this collection, pp. 7-8). I do not think that this distinction commits me to serialism however. In my 2013 book, I propose that “a mind should primarily be seen as consisting of a hierarchy of control-and-monitoring loops, and their essentially dynamic interaction with the world, rather than as constituted by the successive states that emerge from this interaction”. Examples of how the dynamics at lower levels of representation can influence higher levels, and the converse, are discussed in chapters 11 and 12, where the case of schizophrenic delusions is analyzed. Hence, I have no problem with the view that low-level appraisal affects higher-level appraisals: these types of influences are part of what it is to have a hierarchy of control. This does not mean, however, that predictive appraisal and retrodictive appraisal should be conflated: they have a different evaluative function, and are based on different dynamic cues. This does not mean, either, that a concept-based judgment can easily influence an affordance-based appraisal. The difficulty of having a prolonged *strategic control* over one’s feelings (based on what one knows, as in the anagram experiment), originates in the different roles of associative cues and inferential relations between concepts in mental activity.³

Iuliia Pliushch is right, however, when observing that I stick to the distinction between feelings and their propositional re-description. From the viewpoint of action theory, this distinction corresponds to the contrast between reacting and acting strategically. I subscribe also to her remarks on p. 6, according to which goal representations might change affordance-sensings. The point is: how sustained is this change? A conceptual re-description tends to modify one’s representation of the context, and hence of one’s goals, which might either favor or re-

³ This point is developed in Proust (2014). A third form of action, habitual or routine action, is claimed to pertain to a second affordance-based system with its own agentic feelings of opportunity.

duce *further elicitation of feelings* (for example, by being ashamed of having felt anger), and even inhibit the influence of feelings on action. This is the case for the participants' epistemic decisions in phase 2 of the Anagram Experiment discussed in the section 5.2 of the article. Their ability to control their feelings, however, cannot resist time pressure and/or divided attention in phase 3.

On the view that I propose, feelings can only be sustainably modulated by having other feelings replace them. There are both automatic and strategic ways of enhancing one's feelings through other feelings (see [Proust 2014](#)). Feelings can easily be enhanced by enriching the associative representations constituting an affordance. Deliberately suppressing them, or reorienting them to new targets, however, is very difficult (as rejected lovers know all too well). The Confucian moral practices offer a very good example of a strategic attempt to train new moral feelings in followers (see [Reber 2013](#)). As Rolf Reber shows in his fascinating analysis of what he calls critical feelings, strategically redirecting one's feelings to new targets can only be performed by manipulating the fluency of one's own re-descriptions and conceptual rules for acting morally. In other terms, the agents need to be trained until they entertain feelings of ease of processing (i.e., feelings of fluency) when activating target concepts and inferences, rather than merely trying to immediately subsume their own initial feelings under critical concepts.

7 Self-deception and metacognition

Iuliia Pliushch finally makes an interesting suggestion: when self-deception occurs, the believer senses a metacognitive feeling of uneasiness, indicating that her underlying belief-forming process is faulty. This suggestion offers an account of the tension that arises while forming a belief on the basis of motivational, rather than evidential grounds. It would be wrong to interpret her proposal as the claim that finding faulty a belief, or a belief-forming process, involves an appraisal of the content of the belief, or of the kind of process that has been used to form it. As I understand her, Pliushch is rather claim-

ing, as psychologists and neuroscientists of metacognition do, that the mind is able to detect fault in the dynamical properties of the underlying processes. Pliushch argues further that, in contrast (she claims) with my own proposal, monitoring not only occurs "before or after a cognitive process, but also *during it*". There is no real conflict, however, about this claim. Presence of intermediate monitoring depends on the temporal extension of the mental action considered. When confronted with perceptual or memorial uncertainty, there is only control-based, mainly unconscious, intermediate monitoring; intermediate becomes prominent, however, in prolonged, effortful actions, such as problem solving ([Ackerman 2013](#)). I agree with Pliushch, however, that representing a mental action merely in terms of a starting and end points misrepresents the facts: it is based on a serial view that does not fit the dynamic character of metacognition (as already discussed in section 7 above). The evidence presented in [Proust \(2013\)](#) suggests that retrospective evaluation is based on the underlying dynamic of the *whole* action (the rate of accumulation in favor of a dominant response, as well as the dispersion of the neural responses), while predictive evaluation is based on the dynamics elicited by the command for this action, as compared with a stored standard (the complexity of the feedback used is addressed in [Koriat et al. 2006](#)). An epistemic evaluation, however, has two functions: stop the action, and encourage its continued performance, hence the role of polar valence in motivating action, which is reflected in the bi-partition of evaluations in two classes. This is in close agreement with how predictive coding, as any other theory of emotion and action, describes the facts.

Does predictive coding offer *new* insights on metacognition? The concept of "transition probabilities" mentioned by Pliushch, is shared by all theorists working on neural dynamics, as well by theorists of recurrent feedback; the concept of free-energy minimization, related to the minimization of surprise, seems *prima facie* to be consonant with [Rescorla & Wagner's \(1972\)](#) well established model of reinforcement learning. There is an internal connection

between free energy minimization and the evaluation of one's own uncertainty, because it is adaptive to predict one's chances of being incorrect, and hence avoid surprising failures. The concept of free energy, however, is no more equipped to provide any mechanistic account of brain function as any other evolutionary theory. "It is nothing more than the principle of least action applied to information theory", Friston recognizes (Friston et al. 2012). Indeed a prominent problem remains to be solved, concerning how priors vary as a function of task demands and of environmental statistics. Unpacking the principle across adaptive time-scales and survival contexts is indeed a complex future goal. Ways in which predictive coding might enrich the analysis of metacognition with new descriptive, operational tools or new functional explanations remain, then, to be specified.

Pliushch claims further that a first step in the proposed metacognitive theory of self-deception consists in recognizing that metacognitive feelings must be "extended to unconscious belief forming processes". If what is meant is that the dynamic properties that elicit feelings belong to such processes, there is universal agreement on this claim (see the so-called "cross-over principle" between unconscious heuristics and representations (including beliefs) and conscious feelings in Koriat 2000). What is meant, then, by the suggested "extension" is unclear. If what is meant, rather, is that the feelings themselves might be unconscious, this is a possibility that is taken seriously in studies of metaperception in blindsight patients (Reder & Schunn 1996). The very existence of such feelings complicates the phenomenologist's task. A second step is claimed to consist in "clarifying the representational content of tension". Although more detailed work needs to be done in order to better understand the contrast between perceptual and conceptual fluency, intuitivity is generally identified as a variety of what experimental psychologists call "feelings of fluency". One suggestion is that what creates feelings of tension or dysfluency in self-deception is not merely the representation that "the cognitive process violates some important goal representation", but rather, that it violates an implicit heuristic of

self-consistency, as discussed in Koriat (2012). Another suggestion is that tension has to do with the realization that the effort initially planned for a current task needs to be upgraded, which is a source of anxiety (Ackerman 2013). In summary: belief-forming processes are known to elicit metacognitive feelings. It remains to be shown how a metacognitive analysis of self-deception might enlighten philosophical and epistemological views about it. Self-deception is a good test case for making the point that conceptual-inferential processing also conveys non-conceptual information.

8 Serial versus dynamic properties of cognitive processes

As noted in the title of an article by Koriat et al. (2006), the relations between control and monitoring in the production of metacognitive feelings are very "intricate". Iuliia Pliushch's insightful comments have initiated what I hope to be a useful clarification of another aspect of feelings (whether metacognitive or not): their relations with propositional thoughts. Feelings elicited by tasks that are conceptually characterized do not become *ipso facto* conceptually penetrable: this difficult, unintuitive claim is often misunderstood and resisted for wrong reasons, which does not mean that it would resist any reason! The objection related to serialism was odd, given my own interest in the dynamic properties of the mental processes as offering a source of information that stable propositional properties of mental contents cannot provide. Once prediction and post-evaluation are identified as two major functions in metacognition, it is indeed important to emphasize that metacognitive processes of each kind are dynamic, and rely on various types of re-afferent feedback. Epistemic decisions, however, once made, are discontinuous by design, which turns the pre-decisional confidence level into a final evaluation that triggers or inhibits the corresponding action. Hence, a contrast must be maintained between how to select a goal and determine the level of effort needed to achieve it (i.e., a control command), on the one hand, and monitoring progress toward the goal, on the other hand.

Each form of metacognition elicits feelings. This does not mean that the two functions need to be serially executed: for long, effortful tasks, agents need to frequently revise their level of effort and of success expectancy, by monitoring over time their progress through associated heuristics and feelings.

References

- Ackerman, R. (2013). The diminishing criterion model for meta-cognitive regulation of time investment. *Journal of Experimental Psychology: General*, 143 (3), 1349-1368. [10.1037/a0035098](https://doi.org/10.1037/a0035098)
- Carver, C. S. & Scheier, M. F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological review*, 97 (1), 19-35. [10.1037/0033-295X.97.1.19](https://doi.org/10.1037/0033-295X.97.1.19)
- (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological review*, 97 (1), 19-35.
- (2001). *On the self-regulation of behavior*. Cambridge, UK: Cambridge University Press.
- Friston, K. J., Thornton, C. & Clark, A. (2012). Free-energy minimization and the dark-room problem. *Frontiers in Psychology*, 3 (130). [10.3389/fpsyg.2012.00130](https://doi.org/10.3389/fpsyg.2012.00130)
- Koriat, A. (2000). The feeling of knowing: Some metatheoretical Implications for Consciousness and Control. *Consciousness and Cognition*, 9 (2), 149-171. [10.1006/ccog.2000.0433](https://doi.org/10.1006/ccog.2000.0433)
- (2012). The self-consistency model of subjective confidence. *Psychological Review*, 119, 80-113.
- Koriat, A., Ma'ayan, H. & 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: General*, 135 (1), 36-69. [10.1037/0096-3445.135.1.36](https://doi.org/10.1037/0096-3445.135.1.36)
- Pliushch, I. (2015). The extension of the indicator-function of feelings. In T. Metzinger & J. M. Windt (Eds.) *Open MIND*. Frankfurt a. M., GER: MIND Group.
- Prinz, J. (2010). For valence. *Emotion Review*, 2 (1), 5-13. [10.1177/1754073909345546](https://doi.org/10.1177/1754073909345546)
- Proust, J. (2013). *The philosophy of metacognition. Mental agency and self-awareness*. Oxford, UK: Oxford University Press.
- (2014). Time and action: Impulsivity, habit, strategy. *Review of Philosophy and Psychology*. [10.1007/s13164-014-0224-1](https://doi.org/10.1007/s13164-014-0224-1)
- (2015). The representational structure of feelings. In T. Metzinger & J. M. Windt (Eds.) *Open MIND*. Frankfurt a. M., GER: MIND Group.
- Reber, R. (2013). Critical feeling. In C. Unkelbach & R. Greifeneder (Eds.) *The experience of thinking* (pp. 173-189). Hove, UK: Psychology Press.
- Reder, L. M. & Schunn, 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). Mahwah, NJ: Lawrence Erlbaum Ass.
- Rescorla, R. A. & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.) *Classical conditioning II: Current research and theory* (pp. 64-99). New York, NY: Appleton Century Crofts.
- Solomon, R. C. (2003). Against valence ('positive' and 'negative' emotions). In R. C. Solomon (Ed.) *Not passion's slave* (pp. 162-177). Oxford, UK: Oxford University Press.
- Stepper, S. & Strack, F. (1993). Proprioceptive determinants of emotional and nonemotional feelings. *Journal of Personality and Social Psychology*, 64 (2), 211-220. [10.1037/0022-3514.64.2.211](https://doi.org/10.1037/0022-3514.64.2.211)