First-Person Science of Consciousness. Theories, Methods, Applications

May 6-8, 2021

Witten/Herdecke University, Germany

On the predictive and motivating roles of metacognitive experiences

CONSEIL SCIENTIFIQUE DE L'ÉDUCATION NATIONALE

Joëlle Proust
Institut Jean-Nicod, Paris
http://joelleproust.org

Institut | Nicod

Outline

1. Prediction & motivation in cognitive agency
2. Controversy about the concept of metacognition
3. Comparative evidence
4. Why should metacognitive evaluations be conscious?

1. Prediction & motivation in cognitive agency

Defining action

- " Behavior is purposive when its course is subject to adjustments which compensate for the effects of forces which would otherwise interfere with the course of behavior."

- *Harry Frankfurt "The problem of action"*

Defining cognitive action

A cognitive action consists in using **compensatory mechanisms** to maintain a trajectory **toward a cognitive (informational) goal**

Defining cognitive action

This definition captures the interaction between

- Motivation → Goal selection
- Evaluative prediction → progress to the goal, compensations

Examples of cognitive actions

Controlled perception :- trying to discriminate/ to categorize	**Controlled memory** trying to remember, to learn	**Controlled reasoning** trying to solve a problem,to understand an argument, to plan	**Communicate** trying to persuade	**Controlled imagination** trying to invent, to extend one's thinking

The two functions of metacognition

- **CONTROL**

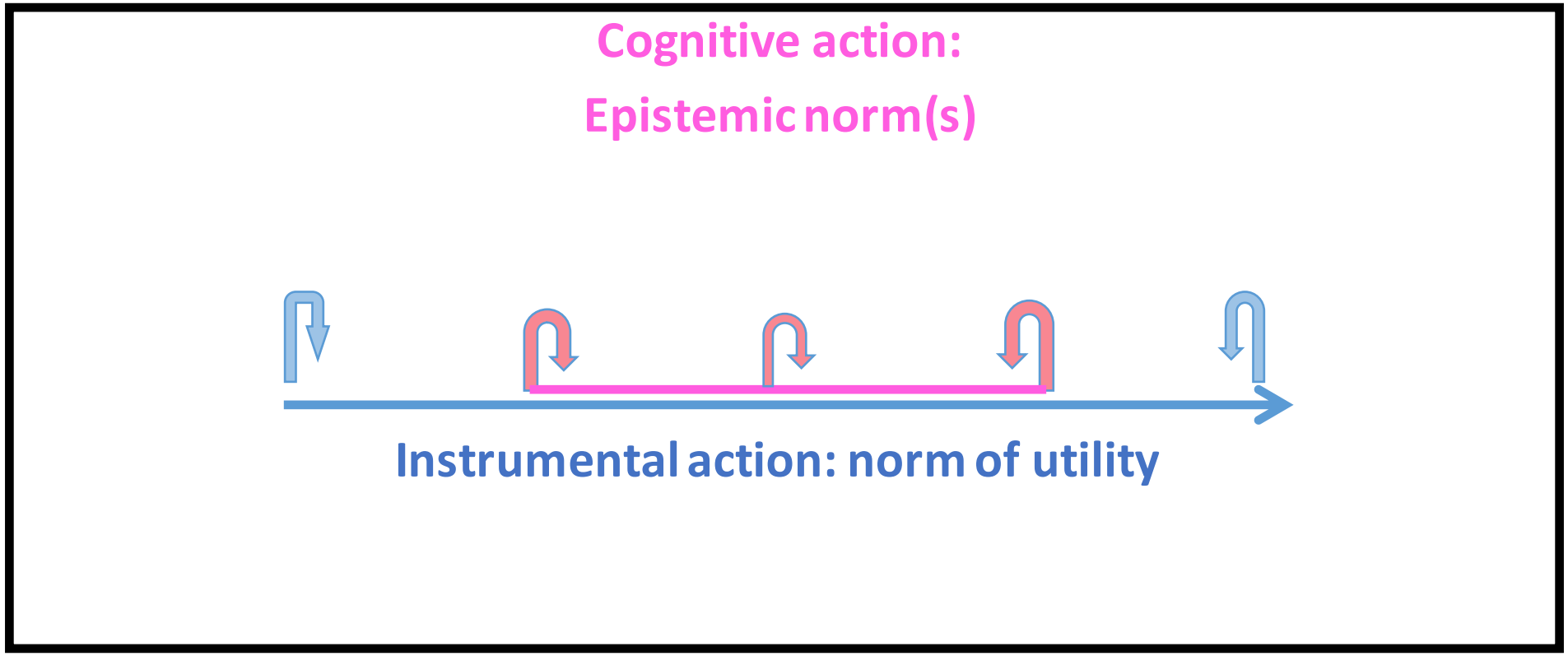
Consists in selecting a cognitive action, and adjusting the amount of effort to perceived demands and expected benefit of the task.

Control is itself nonconscious, but feedback generated by control is the source of specific metacognitive feelings : e.g. sense of effort. (Koriat et al. 2006)

- **MONITORING**

Consists in using conscious feedback from control and from performance (noetic feelings) to predict failures and discrepancies from expected cognitive outcomes

The embedded structure of cognitive action



Predictive /evaluative function of metacognitive feelings

The function of metacognitive feelings consists in evaluating one's own cognitive actions for their

1. **Goals :**
    1. **Feasibility: is this goal achievable at all at present ? (e.g. feeling of knowing)**
    2. **Chances of successful completion (feeling of confidence)**

2. **Progress toward the goal: does the action move in the anticipated direction? (e.g. feelings of understanding/confusion, feelings of error)**

3. **Final outcome : it is correct ? (feeling of being right)**

Goal-related predictive feelings	Process-related evaluative feelings	Result-related evaluative feelings
Feeling of knowing	Feeling of confusion	Feeling of being right/wrong
Tip of the tongue	Feeling of error	Feeling of uncertainty
Feeling of difficulty / ease	Feeling of incoherence	Feeling that one learned
Feeling of familiarity	Feeling of incomprehension	Eureka feeling
	Feeling of interest/boredom	

2. Controversy about the informational sources of metacognition

What is the informational source of metacognitive feelings ?

- **Two views:**

  - **The self-attributive vie**w (Carruthers, 2009, 2011)

  Metacognition requires using concepts of knowledge, memory, etc.

  It depends on mindreading and associated conceptual inferences

  - **The procedural view** (Proust (2007, 2013)

  Metacognition is a prediction that is based on the affective feedback from cognitive actions

  It does not depend on mindreading, although it can be combined with it in humans.

What is the informational source of metacognitive feelings ?

- Controversy has **not** been about **the concept-based** notion of metacognition (attributing to oneself states of uncertainty)`

- Controversy has been about the existence of an **experience-based** type of metacognition.

Main argument in favor of procedural metacognition

- Non-humans (rhesus monkeys, rodents), evaluate their uncertainty in perceptual and memory tasks as reliably as humans do

- Non-humans however, do not have a concept of mind, perception, memory, uncertainty, and are not able to metarepresent their mental states.

- What they can do, however, is predict successful or unsuccessful cognitive outcomes through reliable predictive heuristics.

3. Comparative evidence
Emotion, not judgment primarily involved

Main types of tasks eliciting decisions under uncertainty

- Seeking for information (SI) tasks:

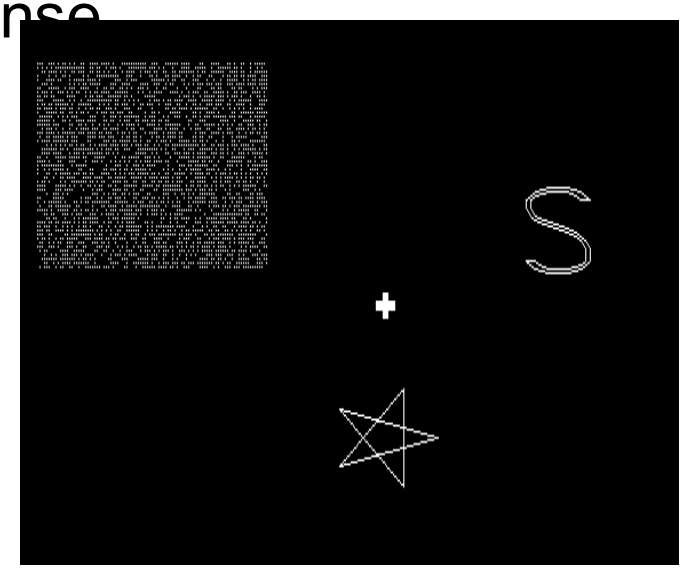
  Will an animal ask for information only when needed ? (Call & Carpenter, 2001)
- Buying hints when learning (Kornell, Son, & Terrace, 2007) .
- Choose-or-decline-to-respond (« opt out ») paradigm (Smith et al., 2006, Beran et al. 2010)
- Wagering tasks (betting on a given response).
- Action control based on internal confidence feedback (Beran et al. 2015)

What monkeys can do

- They can solve the 5 types of tasks listed using visual discrimination and memory paradigms
- They can generalize their uncertainty responses to new tasks.
- They can also use uncertainty responses with blocked feedback (which shows that these are not conditioned responses)

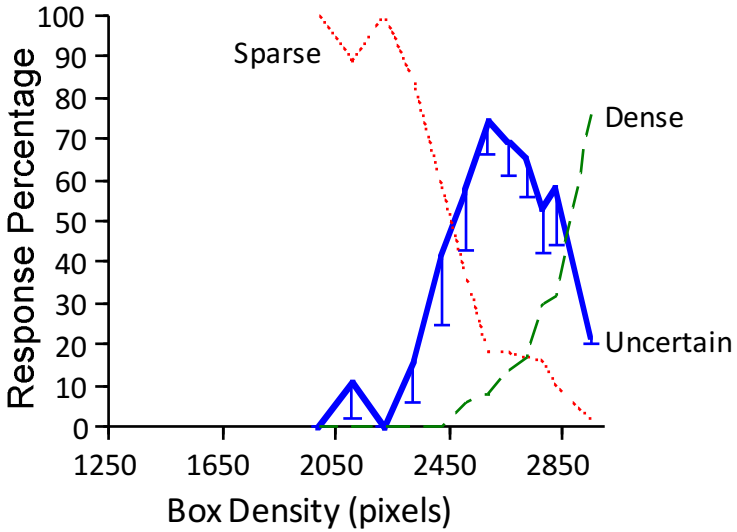
Sparse

Dense

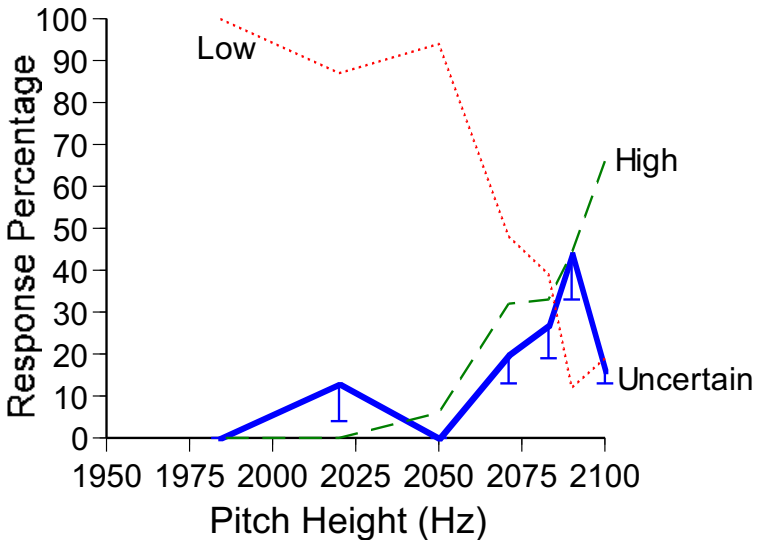


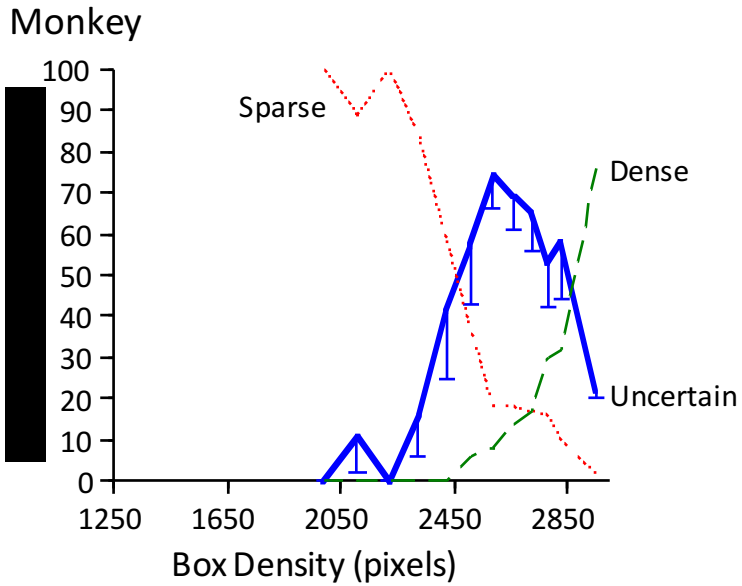
Uncertain

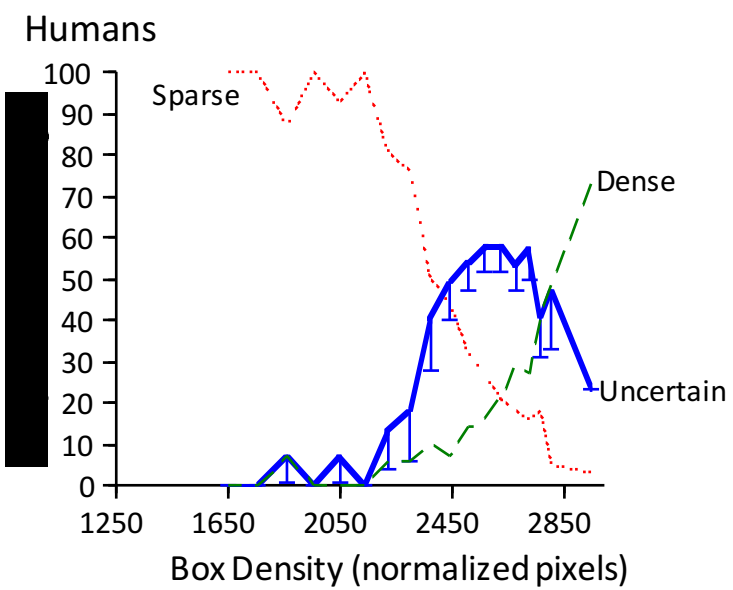
Macaque



Dolphin

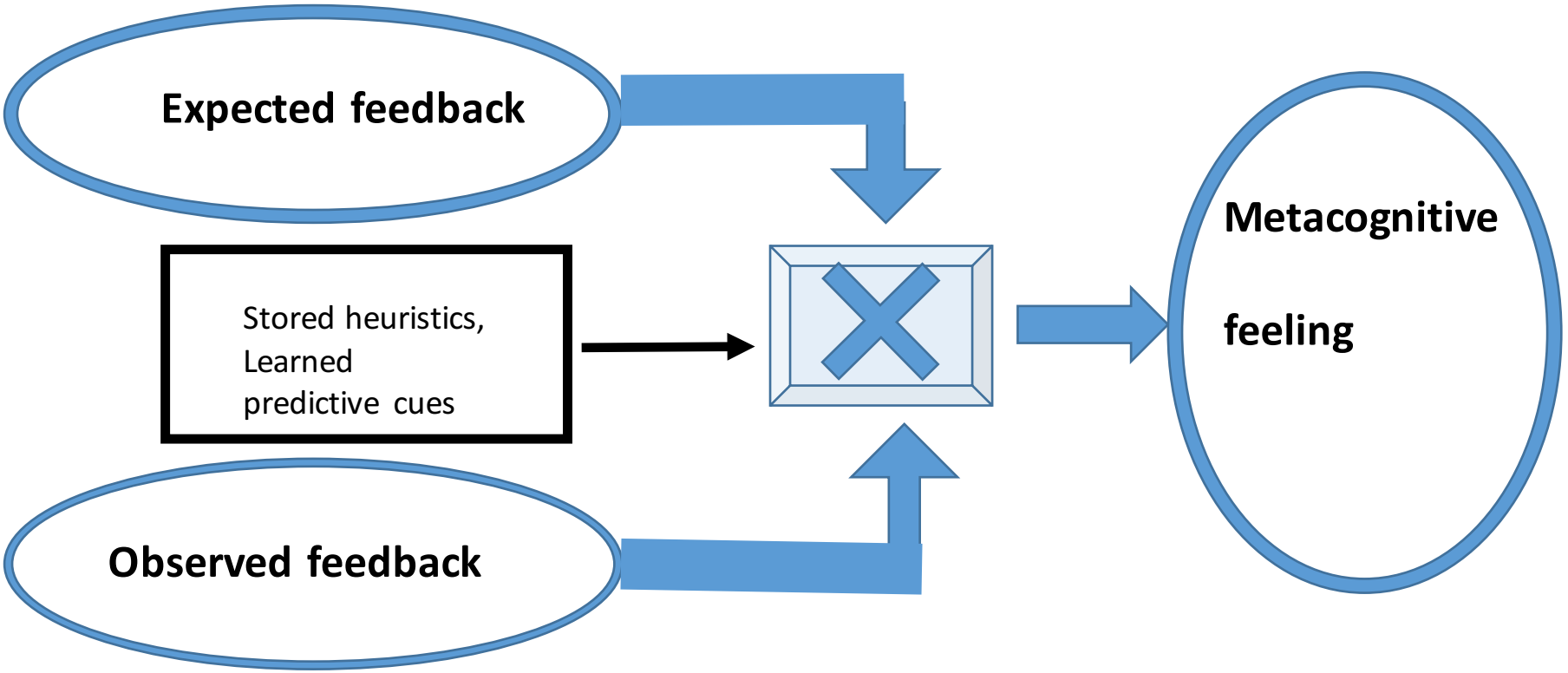



Monkey
100
90
80
70
60
50
40
30
20
10
0
Sparse
Dense
Uncertain
1250
1650
2050
2450
2850
Box Density (pixels)


Humans
100
90
80
70
60
50
40
30
20
10
0
Sparse
Dense
Uncertain
1250
1650
2050
2450
2850
Box Density (normalized pixels)

3. Hypotheses about the informational basis of metacognitive predictions and evaluations

When evaluation is needed, a comparator with the following functional structure



Question 1

- Are the predictive cues used to evaluate one's cognitive performance of a conceptual or non conceptual kind ?

- **Functional claim 1**

Evaluations and predictions of cognitive feasibility, cognitive success or error crucially **depend on nonconceptual cues,** extracted from the on-going cognitive activity in **the underlying vehicle**

Variety of unconscious heuristics

- Cue familiarity: elicited by the terms of the question (Reder, 1987)
- Overall accessibility of pertinent information regarding the target: elicited by the activity triggered by the question (Koriat, 1993)
- Fluency heuristic: responses that come to mind quickly are deemed correct. (see: priming experiments as a source of confidence enhancement: Kelley & Lindsay, 1993)

Recently discovered heuristics

- Predictive neural dynamics: Kepecs & Mainen (2012).
- Predictive interoceptive cues from the respiratory, circulatory, digestive, and endocrine systems: Barrett & Symons (2015), Park & Tallon-Baudry (2014).
- Proprioceptive (postural, facial) cues (corrugator/zygomaticus): Stepper & Strack (1993)

Functional claim 1

- Predictive cues **are automatically collected** during cognitive actions just as they are during bodily actions, **through reinforcement learning,** in order to anticipate for potential trajectory derailments from the action program, and guide compensatory moves when needed.

- Tensions in the joints (e.g. elbow or wrist) predict success in a tennis move (if they coincide with expected feedback)

- They do not **metarepresent** one's (un)certainty in scoring a point.

Proust (2019)

Question 2

Many comparators do not generate any conscious feeling

- Ex: the unconscious error signal driving immediate correction when typewriting (Logan & Crump, 2010)

→ **Why are metacognitive feelings conscious ?**

4. Why should metacognitive feelings be conscious?

Asher Koriat addressed this question in a foundational paper, where he articulated the "cross-over principle"

ELSEVIER

Consciousness and Cognition

Volume 9, Issue 2, June 2000, Pages 149-171

Regular Article

The Feeling of Knowing: Some Metatheoretical Implications for Consciousness and Control ☆

Asher Koriat

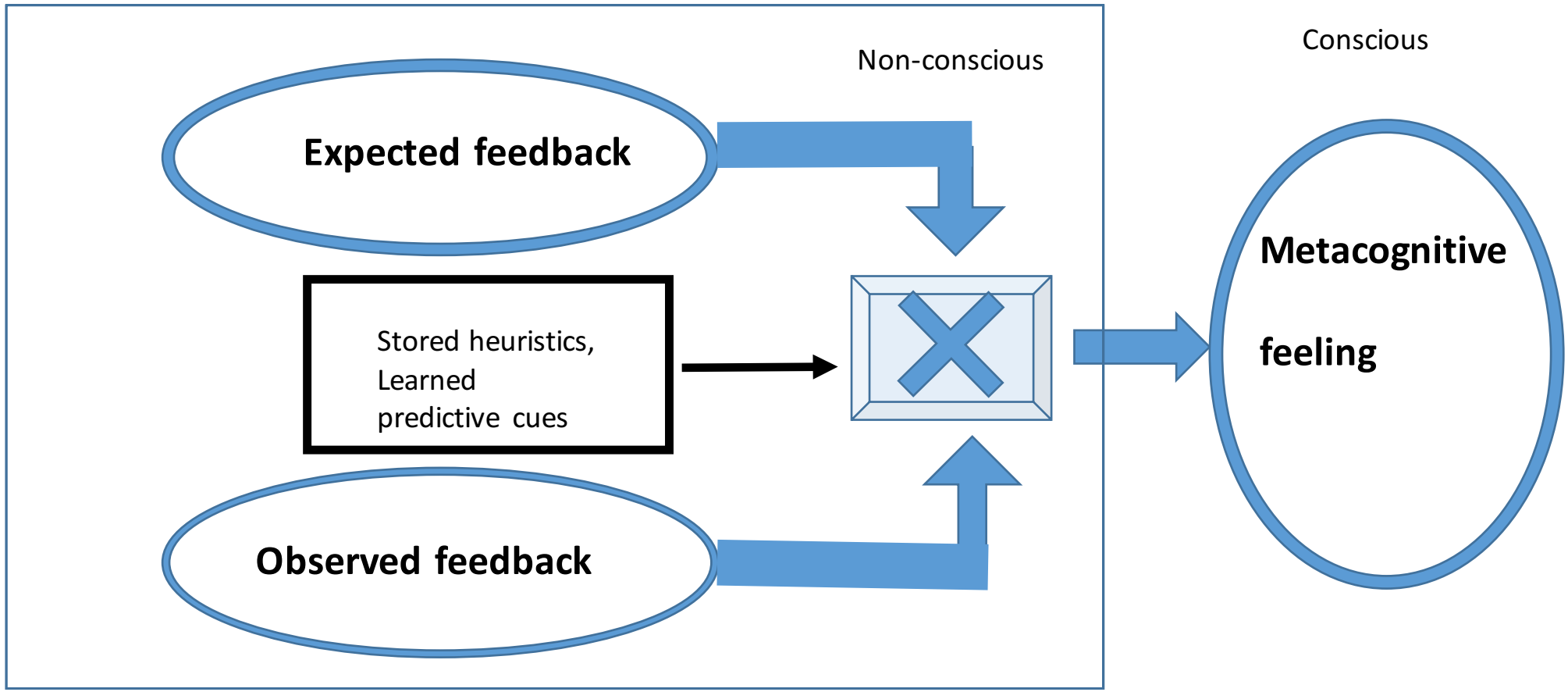
Cross-over model (Koriat, 2000)

Metacognitive experiences are

- **implicit** in their antecedents (unconscious processing)

- (once formed): **explicit** in their consequences (controlled decision-making)

At each point of evaluation, a comparator with the following functional structure



The cross-over principle

Subjective experience of uncertainty mediates

1. **a nonconscious cause (= predictive heuristics)**

2. **a cognitive decision**

→ Feelings are conscious because only conscious states

enable *flexible guidance of decision-making*.

Flexibility has two dimensions:
A. Evaluative flexibility: **as a function of context**

- the capacity of **inhibiting** an impulsive or a routine evaluation when needed requires using an explicit representation for resisting an implicit influence.
- For example, explicit rules (verbally expressed or internally represented) can overcome an overlearned response.

B. Response flexibility as evaluative integration

- Response flexibility may require integrating various evaluations (such as epistemic, instrumental, social etc.)

- There is evidence that
  - Evaluating likely reward (assessing a **world opportunity**)
  - Evaluating likely cognitive success (assessing a **cognitive opportunity**)

  are processed by distinct brain structures. (Kepecs & Mainen, 2012).

B. Response flexibility as evaluative integration

However, **a single decision to act has to be made**, integrating subjective cognitive uncertainty and objective world uncertainty

→ evaluations must share "a common currency" (Sugrue Corrado, & Newsome, 2005).

This common currency might consist in subjective overall desirability, a conscious feeling integrating various valuations.

Elaborating on Koriat's cross-over principle

- Proust (2015,in preparation)

- Koriat observed that the implicit cues relevant to metacognitive feelings are **vehicle cues (such as ease of processing), not content cues (such as inferential relations)**.
- However, metacognitive feelings are always felt as **being about a task**.
- how is this dimension of the subjective experience to be explained ?

The initial cross-over principle between

- **Unconscious vehicle cues** (and associated implicit heuristics), which carry **predictive**/evaluative information
- **Subjective feelings** expressing the **predictive computational output** with respect to a task or cognitive opportunity

Cross-over principle

Beyond the cross-over principle

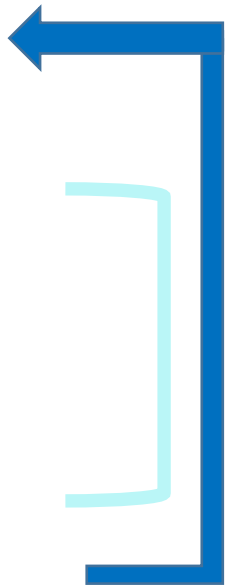
- **Are there conscious** cues identifying **the type of ongoing cognitive task at a given time t?**
- **Unconscious predictive vehicle cues** (and associated implicit heuristics), **which carry predictive**/evaluative information
- **Subjective feelings** expressing the **predictive computational output** with respect to a given cognitive opportunity

Question unaddressed

Cross-over principle

Beyond the cross-over principle

- **Are there conscious** cues identifying **the ongoing cognitive task at a given time t?**
- **Unconscious predictive vehicle cues** (and associated implicit heuristics), which carry the basic information about the cognitive affordance, its
- **Subjective feelings** expressing the **predictive computational output** with respect to a task



Unexplained: How are conscious noetic feelings **felt as being about the ongoing task?**

How is the feeling felt as being about this task ? The hypothesis of cognitive affordance sensings

- Subjective experiences are much more than a mere signal for guiding behavior.
- They **have their own representational structure**, associating

➢ Conscious elements, acting as indexical pointers to
  - world or cognitive opportunities or risks
  - Action programs

➢ Non conscious elements
  - Associative cues collected over time by reinforcement learning
  - The details of action programs
  - The selection of the most relevant affordance sensing in a context

What is an affordance sensing ?

- A type of **nonconceptual representation** that predates conceptual propositional attitudes
- This representation has an **evaluative-predictive function** rather than a descriptive function.
- Its structure is based on **two kinds** of affective gradiency that spread over to the associated predictive cues
- Its function consists in selecting and launching the action program appropriate to the current context.

- On affordance sensings, (under various names): see Bermúdez (2003), Campbell (1993), Cussins (1992), Dreyfus & Kelly (2007), Dummett (1993), Gawronski & Bodenhausen (2006), Gendler (2008), Griffiths & Scarantino (2009), Nanay (2013), B.C. Smith (1996), Strawson (1959).

The **conscious core** of this representational structure

- A motivational **goal-index** (the specific nature of the opportunity or risk is captured in an idiosyncratic image)
- **An evaluative prediction:**
  - **valence** + or - : *motivational gradient scale* (preparing action type)
  - **intensity**: *intensity gradient scale* (assessing importance and effort regulation)
- **Embodied in *simulating markers* produced at time t**
  - **express reactive affect, hence help select an action type**
  - **Distributed in various bodily "affective metrics"** (conscious somatic markers)

Indexing a goal or a task through a new cross over:

Vehicle → **task**

- **phonological-auditory experiences,**
- **visual imagery**
- **proprioceptive imagery**
- **Visuo-motor imagery**

Enabling function of conscious indexing:

Goal-monitoring of:

- **Conceptual reasoning tasks, including planning**
- **Mathematical tasks, spatial reasoning**
- **Meditation**
- **Sport training, drawing**

The nonconscious elements associated to the conscious representational core

- Depend on reinforcement and social learning (in humans):
- Predictive cues collected over time in relation to a given cognitive affordance, for example:
  - Context-sensitive self-representation
  - Types of cognitive goals that are "for me" or "not for me"
  - Naive implicit beliefs and theories about the association between gender or social origin and cognitive abilities
  - Implicit theories of mental competences
  - Implicit theories about cognitive task demands

Summary of the present proposal

1. Conscious feelings in the narrow sense of affective signals belong to **a larger representational structure, called affordance sensing.**
2. **Affordance sensings occur at the evaluative junctures of cognitive actions.**
3. **Because they include an action program motivated by felt** valence and intensity, they are experienced **as being about the indexed task (goal, situation) not about the somatic markers of their affective value.**
4. They are **entertained even when no reference to mental concepts is available**. Concept-based simulations, when available, **can be associated** to their conscious core.


THANKS FOR YOUR ATTENTION
DOWNLOAD THIS PRESENTATION AND RELATED ARTICLES : http://joelleproust.org

Structure of a cognitive affordance sensing
(enabling nonconceptual metacognitive evaluation)

- **Type of task: conscious goal-indexing (e.g. "name of __?)**
- **Conscious evaluation of task progress (noetic feelings):**
  - **valence** + or - : *motivational gradient scale*
  - **intensity**: *intensity gradient scale*
  - **On various bodily "affective metrics"** (conscious somatic markers)
- **time**: *present*
- **Task Location** *Here*
- **Disposition to act**: *persisting, amplifying, opting out*

Proust (2013, 2015, 2016)