Workshop Metacognition

Birkbeck University of London, March 23, 2018

Metacognitive valence in context

Joëlle Proust

Institut Jean-Nicod, ENS, Paris

http://joelleproust.org

Institut | Nicod

2 dimensions in emotion awareness

- **Valence**: hedonic character (pleasant or unpleasant) of a stimulus
- **Arousal:** intensity of the related activity (low or high) subserved by
  - the ascending reticular activating system regulating wakefulness,
  - the autonomic nervous system,
  - the endocrine system.

Valence: definition

- Kurt Lewin used "Aufforderungcharakter" to designate **the attraction or repulsion exerted by a given stimulus on an organism's behaviour,** considered independently from the stimulus identity.

- Tolman translated this term as "valence".

Three reasons for focusing on valence

1. This concept is still not yet stabilized in theoretical psychology and in the philosophy of mind
2. **The concept of valence lays at the intersection of several puzzles in metacognitive studies.**
3. Understanding the *qualitative character* of valence is crucial for analysing the components of conscious experience, in particular the experience of thinking (cognitive phenomenology)

Outline

1. Puzzles about metacognitive valence
2. Solomon and Prinz: "against" and "for" valence: difficulties
3. An alternative view "for" valence
   a. Valence and cognitive goals
   b. Multiple goals and indexing
   c. Projecting valence
4. Objections and responses

1 – Puzzles about metacognitive valence

Puzzles about metacognitive valence

1. **There is disagreement about the affective character of the metacognitive properties of our thoughts**

For Descartes and Locke, properties of ideas such as "clear", "distinct", "evident", "certain", are affect-independent.

Hume observed, however, that

- "the vivacity of the idea gives pleasure"
- Its certainty "prevents uneasiness by fixing one particular idea in the mind, and keeping it from wavering the mind"

Puzzles about metacognitive valence

Some contemporary psychologists claim that metacognitive feelings are "cold", just as the proprioceptive feeling of having one's right arm extended.

Stepper & Strack (1993):

- Effort, familiarity, surprise, hunger, fatigue and the "feeling of knowing" are examples of feelings that have no fixed valence or are evaluatively neutral. These experiences are therefore not considered emotions [by the authors].

- **No analysis provided for these exceptions**

Puzzles about metacognitive valence

- **Others take valence to be an emotional feature intrinsic to metacognitive experiences**
- Zajonc (1968): The exposure effect increases perceptual fluency, and makes an initially neutral stimulus **feel pleasant**
- Activation of facial muscles provides evidence for fluency-generated **pleasure** (Reber et al. 2004, Winkielman & Cacioppo 2001)

- See also: Schwarz & Clore (2007)

Puzzles about metacognitive valence

2. Three puzzles are created by the assumption that metacognitive feelings do have an emotional valence

2a - how does this assumption account for the fact that feelings of fluency **feel differently** across domains ?

2b – why is fluency sometimes **unpleasantly boring**?

2c – Why can the effect of fluency be suppressed if it is found incidental, i.e., non-diagnostic for task-success?

…/

2a - How does a fixed valence view account for the fact that feelings of fluency differ across domains ?

- When watching a painting, a landscape: feeling of beauty
- When reading a text: feeling of beauty, coherence, understanding, truth
- When encountering a known person: feeling of familiarity
- When hearing a repeated or familiar utterance: feeling of truth

- Alter, Adam L., and Daniel M. Oppenheimer. "Uniting the tribes of fluency to form a metacognitive nation." *Personality and social psychology review* 13.3 (2009): 219-235.

Ater & Oppenheimer: Valence is created by fluency + x

Fluency can occur in different proximal cognitive systems:

- semantic priming induces **conceptual fluency,**
- visual clarity induces **perceptual fluency**,
- phonological simplicity induces **linguistic fluency**.

- **Still same influence** on judgments, **independently of how it is generated (and felt) : accepting/rejecting**

Alter and Oppenheimer, 2009

- **Again: a description is offered, not a justification**

2b – Why is fluency sometimes unpleasantly boring?

Response from psychologists:

- Because fluency can in this case **be attributed** to over-exposure.
- Objection: very young infants use boredom as a cue for exploring other aspects of their environment (presumably with no attribution)

Other response :

- because fluency enters a trade off with informativeness. Proust (2016)
- **Still the connection between heuristic trade-off and valence needs to be articulated.**

2c - Incidental versus diagnostic feeling induction

- A cue is **diagnostic** if its valence and motivational effect are actually **generated** by a **genuine epistemically predictive** feature of the task.
  - For example, a feeling of knowing is diagnostic if it is elicited by a cue or set of cues whose function is to predict memory retrieval.

- A cue is **incidental** if its valence and motivational effect are actually generated by a contextual feature that is **unrelated** to the epistemic task
  - For example handwriting discriminability incidentally influences truth evaluation

Psychologists' explanation

- People tend to **attribute** their feelings to the present target of their attention **("naive theory")**

- For this reason, people are easily misled **into interpreting** an incidental feeling of fluency as being task-relevant

(Schwarz & Clore, 2007)

What remains unexplained

2a: "+ x" view: **unclear whether and how** valence **decomposes** into specific character and general disposition to act

2b: In some contexts, stimulus A elicits a feeling of interest that motivates effort to attain an epistemic goal; in others, A elicits rejection. **Why?**

2c: Is a **naive theory** involved in sensing a feeling as related to the occurrent task?

2 – Solomon (2003) and Prinz (2010) about valence

+ x: Robert Solomon's "against valence" (2003)

- Valence-based theories of emotion (VTE) are unable to characterize emotions because they reduce "the complexity of emotion" "to a simplistic and superficial feature of **some emotions in some contexts**" (2003, 5).

- VTE take pleasure and pain to be quantifiable dimensions on a continuum. However,
  - emotions do not **literally involve pain**
  - Pleasures and pains **cannot be compared**
  - "**Suffering and pain** are not two of a kind" (2003,8)
  - Suffering itself is **not a kind of experience.**

Robert Solomon's objections to valence

General critique:

No unique dimension extending from pain to pleasure

- In particular: no "emotional opposites" (p. 16)
- No single valence for most emotions across contexts
- A positive emotion can be negatively toned, and a negative emotion positively toned.

→ **replace valence with judgment-based motivation**

Jessie Prinz's defense of valence

- True that valence does not work in the same way across emotions
- **Because they are embodied, "there is little reason to think that pretheoretically negative emotions have a phenomenological common denominator.**
- P. 8 Despite this conclusion, the feeling theory of valence "derives from a platitude that must ultimately be explained".
- Key idea: **emotions contain both embodied appraisals and valence markers.**
- **Valence markers = inner signals (reinforcers) for behavioral inhibition or activation.**

Jessie Prinz's defense of valence

- Negative emotions are emotions that contain a component that serves as an inner punishment—a kind of signal that says, "Less of this!"
- Positive emotions are those that contain a component that serves as an inner reward—a signal that says, "More of this!"
- Good feelings are just those that carry a reward signal and bad feelings are just those that carry a punishment signal.
- Fear contains a component that signals "less of this" and **thus implicitly represents its own badness**. It is as if emotions are self-appraising, and hence they constitute (self-referential) meta-appraisals.

Where Prinz (2010) meets metacognition

- Valence belongs to goal regulation
- Nonconceptual character of valence
- Embodiment as a precondition for a noetic feeling with a given valence to be experienced

Where Prinz (2010) does not meet metacognition

Reward does not apply in the same way to epistemic goals and to world goals: valence, then, must be related to goal structure, not to goal contents.

when one feels doubt about P, one is conscious of a proposition being doubtful, not of having one's *corrugator supercilii* contract.

Is valence consciously felt or not?

- "In saying that reward and punishment markers are inner states, I don't mean to imply that they are feelings, like pleasure and pain. **I think they are not consciously felt**." (Prinz, 2010, p. 11)

- If valence **is not** consciously felt, then the flexibility of epistemic decisions fails to be explained (Koriat, 2000)

3 – An alternative view "for" valence

Preview

- **What are the determinants of evaluative attitudes that, together, explain the diversity in metacognitive experiences with an apparently similar "abstract" valence ?**
- **Proposed solution**
  A. Valence is inherently **comparative and goal-related**
  B. Subgoals have **each** their **own valence appraisal. Therefore,** a **goal index** is needed to **intentionally bind** the subgoals with their respective expected versus observed values
  C. Projection is needed to intentionally experience the relevant evaluative features as belonging to the outcome

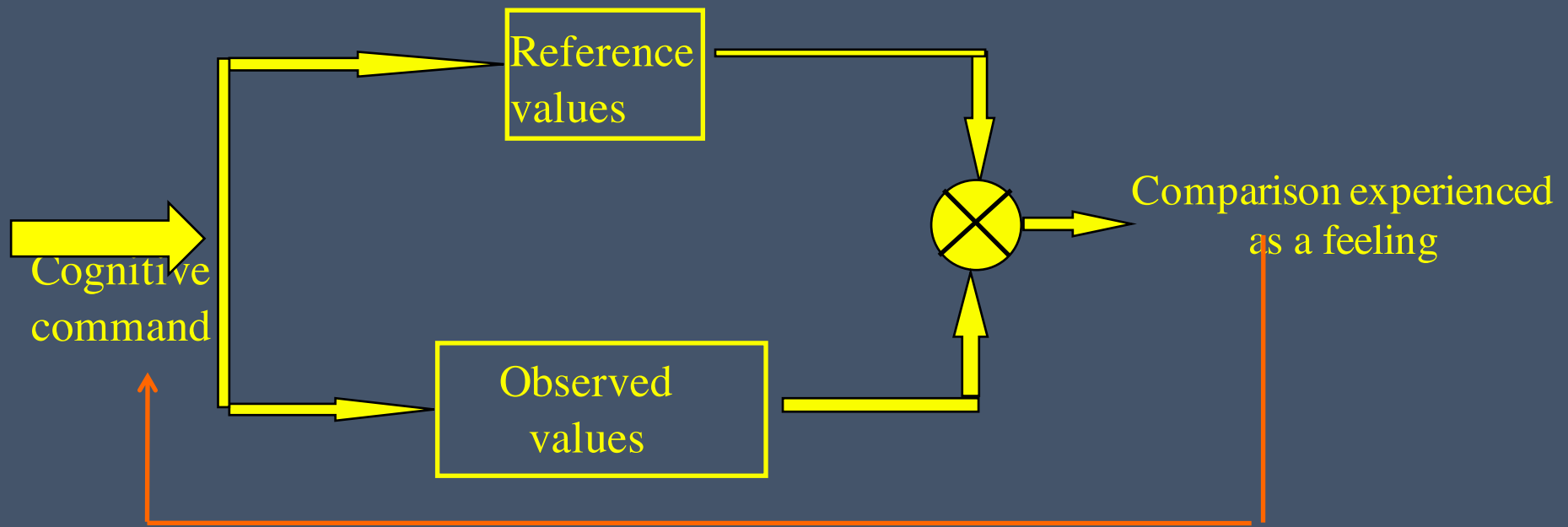
A – Metacognitive valence and cognitive goals

Valence and goal:Carver & Scheier (1990)

- Valence depends not on a goal being attained, but **on how** it is attained **relative to what was expected**
- When people move (physically or psychologically) toward their goals, **they are sensitive to a negative (*i.e.discrepancy reducing*) feedback loop.**
- A sensed value is compared with a dynamic reference value or standard, and adjustments are made, if necessary, to shift the sensed value in the direction of the standard in next attempts.
- Metacognitive control also depends on negative feedback ../

e

Metacognitive feedback loop



How does valence (discrepancy reduction) work?

When a deficit of skill, knowledge, or effort is

- **predicted from task features (including past feedback)**
- **observed from outcome,**

- Appropriate commands range:
  - **more cognitive effort (positive valence)**
  - **Disengaging from the attempt (negative valence)**
  - **Collecting more cues for deciding: this reading looks interesting but might be too hard for me (mixed valence)**

Valence depends on comparing expected/observed values

Table 1

Three Conditions of Behavior Over Time, How They Would Be Construed at the Level of the Action Loop, How They Would Be Construed at the Level of the Meta-Monitoring Loop, and the Affect That Theoretically Would Be Experienced

Depiction of behavior	Action-loop construal	Meta-loop construal	Affect
1. Progress toward goal, at a rate equal to the standard	Discrepancy reduction	No discrepancy	None
2. Progress toward goal, at a rate lower than the standard	Discrepancy reduction	Discrepancy	Negative
3. Progress toward goal, at a rate higher than the standard	Discrepancy reduction	Positive discrepancy	Positive

Habit-based fluency feeling

Uncertainty→ self-doubt → prevention focus

Certainty → self-confidence → promotion focus

Valence serves multiple reference values, not only reward

- A reward-based system would work on **expected gain or loss as part of a single goal-directed action program**.

- **Valence applies to EACH of the many goals and subgoals that people entertain at any given time.**
- For example:
  - Producing true, relevant and coherent statements
  - Producing persuasive statements
  - Affirming one's contextually dominant self-identity and associated values

YAN & OYSERMAN (2018)

Depending on context

Subgoals and associated valences **deeply influence metacognitive control**

- various self-images can be **contextually primed** (independent/dependent, self-efficacious/helpless, promotion/prevention focused)

- Changes in self-representation influence the valence of the associated metacognitive experiences and motivation to act.
  - **This [hard task in maths] is for me! BY Epeer 1**
  - **This [same task] is not for me! BY Epeer 2**

  → **opposed valence in two epistemic peers, with a different self-concept!**

  This solves puzzle 2b

Valence as a "common currency"

- In the case of metacognition (as elsewhere), valence needs to **integrate various motivations** for moving toward or away from a goal for the purpose of making a single decision such as **pursue, revise or stop cognitive goal**!

- Hence, a metacognitive experience at any given time needs to integrate the valences of several subgoals, such as:
  - Correctly completing a task (granting a given ratio of success probability)
  - in a way that supports (rather that threatens) the agent's self-identity as a competent knower.
  - Within a predetermined time/effort boundary

B – multiple subgoals, goal-indexing and valence

Reminder: What was left unexplained

2a: "+ x" view: compatible with the fact that valence is just the outcome of a number of **judgments, not an affective property**

2b:. In some contexts, stimulus A elicits a feeling of interest that **motivates effort** to attain an epistemic goal; in others, A elicits **rejection**. **Why?** **SOLVED : CONTEXT MAKES CERTAIN GOALS OR SUBGOALS MORE SALIENT**

2c: is a **naive theory** responsible for incidental metacognitive cues?

In order to address 2a and 2c

- **We need to understand how task representation is contextually constructed in a way that "interprets" valence in terms of currently active goals.**

- **The naive theory view: people just believe that what feels pleasant is related to goal (Schwarz & Clore) (fails to explain how monkeys use their feelings of knowing).**

- **An alternative : a feedforward model of action explains the predictive binding between perceptions & feelings and current goal.**

A predictive model of cognitive actions

dispenses with theory-building: **subpersonal associative heuristics** are used, not an interpretation of feeling-task relations

- But raises the question of
  - **why specific cues are falsely taken to be goal-predictive** i.e., why a goal with its subgoals (a feedforward model of action) sometimes includes irrelevant cues
  - How valence is experienced **as a feature of the occurrent task**

Proposal: Goal-indexing theory

- A conscious goal representation needs to be kept in working memory until completion to efficiently drive behavior (Baddeley, 2007)
- *Proposal: inclusive goals are indexically represented in working memory.* Example of a search-index:
  - internal speech ["my keys?"]
  - visual icons [of the keys]
- These indexes initiate and sustain the action that they are *indexing*.

Proust, in preparation

Goal-indexing: character

- An indexical is a linguistic expression whose reference can shift from context to context.
- The character of G-indexing is a function on contexts, whose value at any context is the action selected in this context by the agent experiencing it.
  - If the action consists in trying to remember the name of O, the value of the goal index is this action itself (e.g. a piece of inner speech "What's her name?")

Goal-indexing character

- Token-reflexivity: A token of sensory imagery has an adaptive relation to its referent – a specific directive goal – because this token allows the nonconsciously selected goal to be consciously entertained and pursued.

- Idiosyncratic: no *recurrent sensory content* across goal-indexes and across agents, (except through top-down linguistic/conventional influences)
  - Each goal-index is a **phenomenological component of an affordance-sensing**, which is both context- and agent-specific.
  - Its function is to refer to a goal, **thus motivating action and maintaining attention focused on goal until completion**.
  - This may be why goals are only communicated in verbal (e.g. questions) or conventionally **standardized terms** (e.g. traffic lights)

Goal indexing as a precursor & component of other kinds of indexicals

1. Linguistic indexing is based on a token-reflexive rule common to people in a linguistic group.
2. Conventional iconic indexing is based on a mutually recognized token-reflexive rule rule for behaviour coordination (traffic lights).
3. Goal-indexing is a token-reflexive phenomenal marker that idiosyncratically refers to the action to be performed, thus motivating it, eliciting it, and controlling its execution.

- In contrast with 1 and 2: conceptualization of G-Indexing is uneasy and interferes with performance. (tennis service)
- 3 is involved in all forms of ostensive signals.

Goal indexing theory summarized

- G-indexes are **indexes** because only **some of the information** they carry is phenomenologically available (**through goal-relevant fragments of articulatory or visual properties)**.
- Their meaning is offered by **the whole executive structure** that they have as their function to represent.

- **G indexing is a cognitive precursor of demonstratives.**

C – From indexing to projecting

Why is it that

One and the same fluency cue produces different feelings, e.g.:

- **conceptual fluency → understanding reasons**
- **perceptual fluency → familiarity**
- **Linguistic fluency → understanding speech**

Projection as a representational mechanism

- One way of coming to terms with the intentionality of the subjective experience of colour or taste (or other secondary qualities) is to claim that it is *projected* by the perceptual system onto the object.

- Colours or taste arguably belong to a perceiver's experience, not to the object perceived.

- ***Projection*** **automatically associates a given subjective experience with an external, world difference.**

Literal projectivism

- Is the view that our secondary qualitative experiences **are projected to what causes them**: for example redness onto the object experienced as red

- Unattractive consequence: much of our perceptual experience is a category-mistake, because the secondary quality that is projected actually only belongs to the experience, not to the object.

Shoemaker (1996)

Functional projectivism

A given colour experience is functionally projectable to an external object iff such projection

- **enhances the ability to detect** primary object properties such as shape and volume,
- **generates a phenomenally unified** representation of a scene;
- **guides and motivates** proper action selection.

Functional projectivism

Applied to noetic feelings:

A given noetic feeling is functionally projectable to the content of a cognitive task iff:

- It allows an agent to **reliably predict and evaluate** the actual outcomes of the action being monitored.
- It generates a **phenomenally unified normative representation** of a task .
- It **guides and motivates** further action selection.

Calibration is the basis for metacognitive validity

The informational source of calibration mechanisms is both

- **ontologically dependent on embodied feelings** (for their causal mediating role),

- and epistemologically independent because **information trains affects** (not conversely).

General Summary

- Goal-indexing is the experience that represents one's current action as having a specific goal **and** helps sustain execution of the forward model of action that is indexed.
- A noetic feeling is the experience that predictively or retrospectively represents a sequence of an epistemic episode as the rate of moving to or away from this goal.
- Goal-indexing allows the composition of valence gradients in subgoals into a single resultant value.
- Pleasure generated by proximal fluency is experienced as caused by distal goal

In summary

- **Goal indexes for a given task (e.g., reading, watching a picture)**

→ integrate different predictive cues

→ make valence a common currency for decision-making

→ Allow projection of total valence into the currently activated distal target representation (e.g. interesting story, beautiful landscape, etc.)

Objection 1: what role for beliefs ?

- **Do we not crucially need beliefs to plan our actions?**
- **Response**
- Beliefs are a descriptive precondition for selecting goals and planning cognitive actions, but
  - not involved in organizing cognitive action and selection of internal feedback
  - Not needed for monitoring cognitive action

  These functions are better served by affordance-based evaluations & predictions (Proust, 2015, 2016)

Objections 2: why should goal indexing be conscious?

- Just as metacognitive valence **needs to be** consciously felt in order to allow flexibility in epistemic decisions (Koriat, 2000)

- Goal indexing needs to be conscious to efficiently control behavior: defeat interference, explain and justify ongoing behavior.

Objection 3: ad hoc suppositions?

- Are not goal indexing and projection ad hoc mechanisms supposed to solve the puzzles, rather than mechanisms independently supported?

- There are strong independent arguments in favor of these two mechanisms.

- **../**

Functional coherence indexing-monitoring

Research on **"efference copying"** covers the subject that we call "goal indexing".

The causal-intentional relation between efference-copying and comparators has been defended in a number of research areas (from machine learning, to the neuroscience of action and cognitive pathology.

Objection 4: Projection fosters subjectivism about metacognitive feelings

- In the absence of a mediating mechanism ensuring the **informational alignment** of predicted with observed epistemic preconditions and outcomes, projectivism would amount to denying any epistemic value to noetic feelings. **(as a matter of taste and personal convenience)**
- Calibration mechanisms, however, allow agents to revise their decision thresholds in order to **approximate objective norms of evaluation**

Objections

- availability of normative concepts in a formal language cannot be taken as a requirement for having an epistemic warrant in the perceptual case (Burge, 2003),

→ no reason for denying that noetic feelings not only cause behaviour, but constitute a motivation to act in accordance with epistemic norms and desiderata.

Thanks for your attention

This presentation can be downloaded on:
http://joelleproust.org