

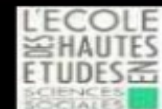
DUKE UNIVERSITY

SSNAP 2023

METACOGNITION
from science to philosophy

Joëlle Proust
Institut Jean-Nicod
Paris

Institut | Nicod



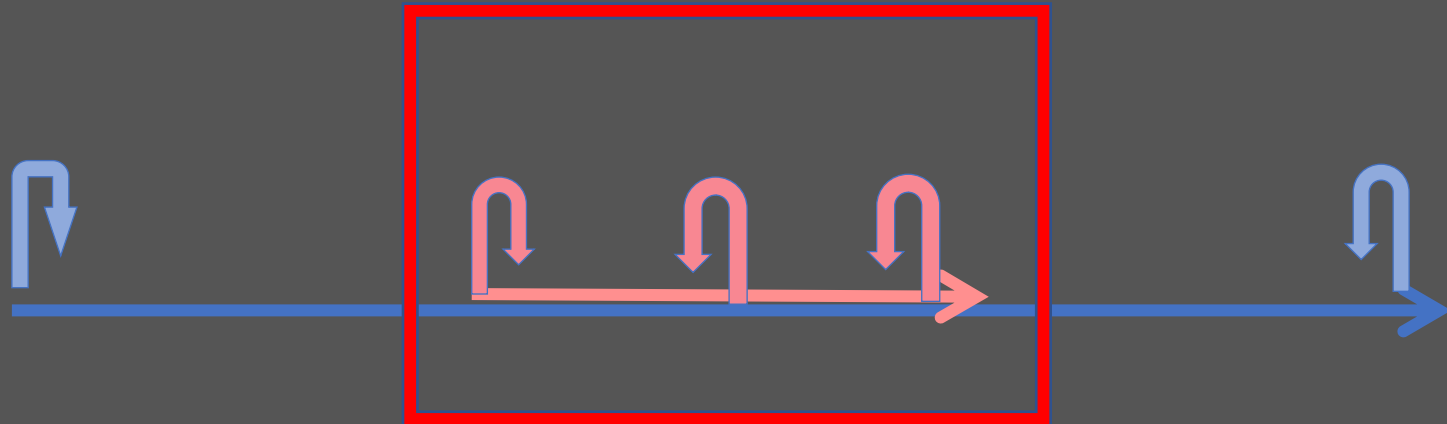
Introduction

What is metacognition ?

- Not "cognition about cognition", which suggests that metacognition (only) consists in representing the **contents** of first-order cognition.
- Rather: it is the ability to predictively or retrospectively assess one's own likely success in a **cognitive action**, such as: trying to remember, to discriminate, to learn, to solve a problem.

Metacognitive evaluations are functional constituents in cognitive actions

Cognitive action



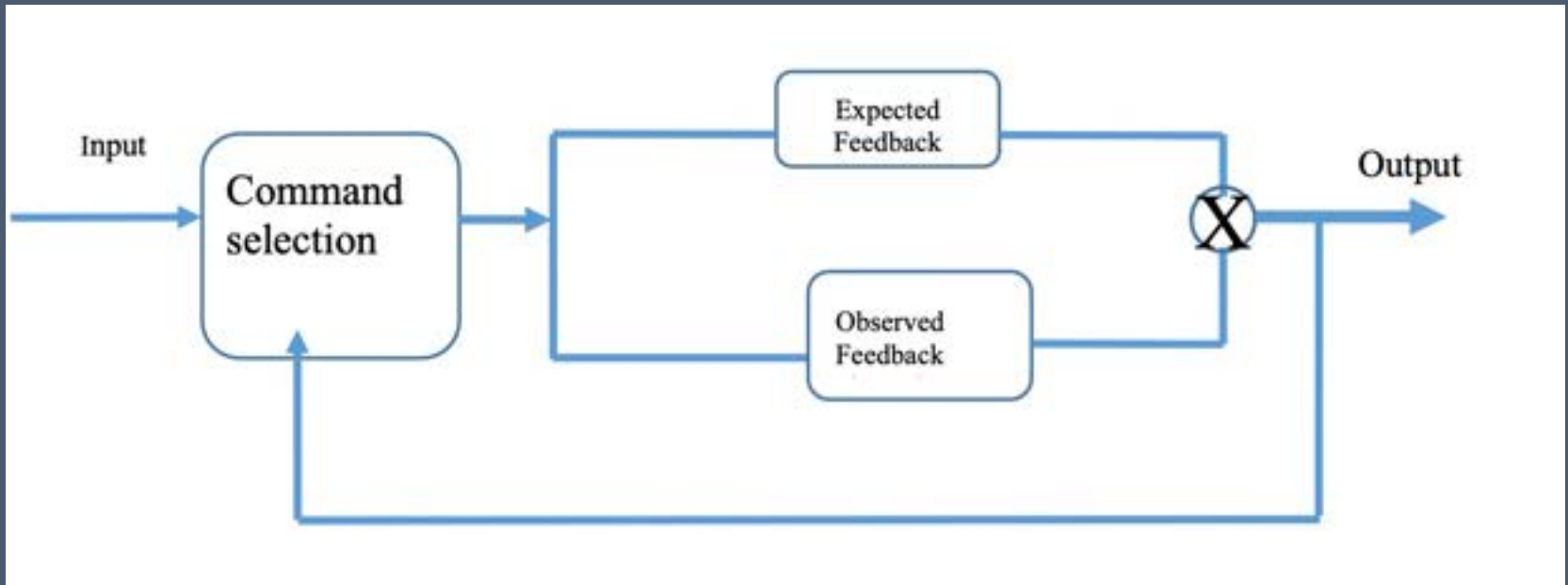
Instrumental action

Two types of mechanisms

- **Procedural metacognition (PM)** is based on nonconscious predictive heuristics and comparators and their affective outputs (metacognitive feelings).
- **Explicit metacognition (EM)** reassesses and occasionally blocks procedural metacognition. Its evaluations are based on episodic memories, beliefs and theories about the on-going task and own ability in this task.

Koriat & Levy-Sadot (1999).

Homeostatic structure of action regulation: procedural

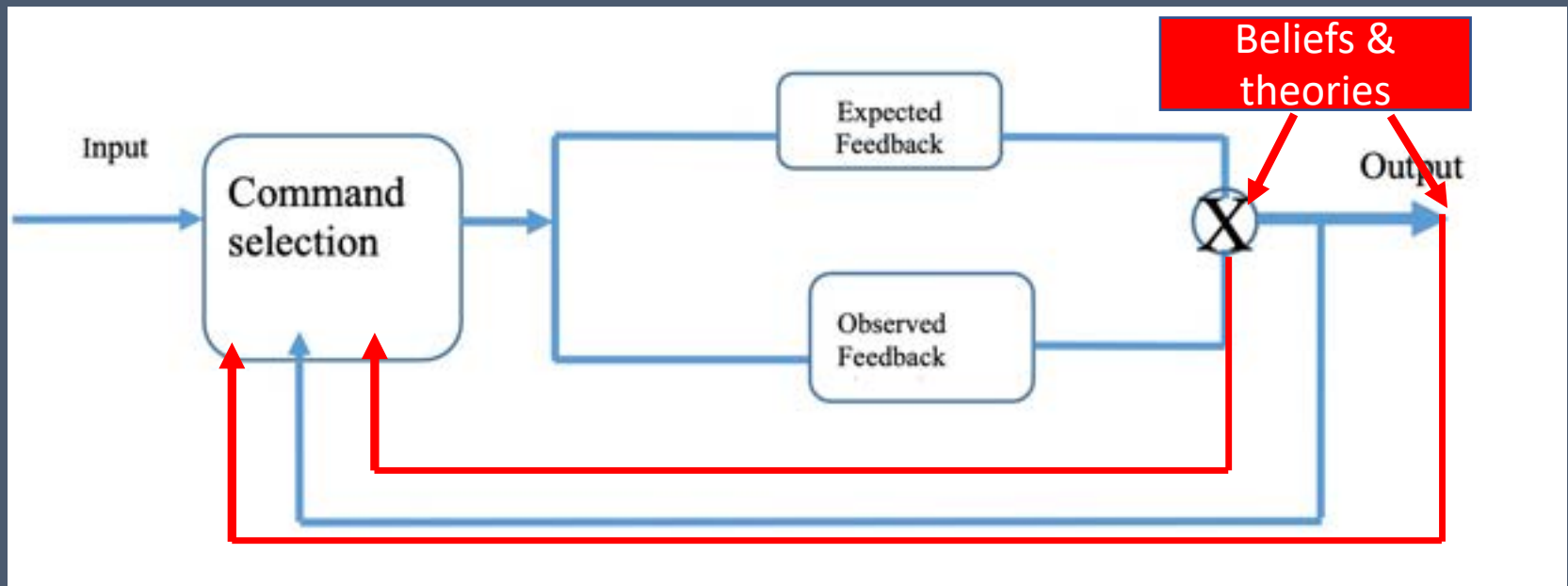


An homeostatic structure based on feelings

has a **dual function**:

- Predicting uncertainty level through feelings with a graded positive or negative valence
- Motivating decision-making based on valence

Homeostatic structure of action regulation: explicit



An homeostatic structure based on beliefs

has a **sociocognitive** function:

- Resisting impulsive decision-making
- Offering verbal justifications for one's decisions under uncertainty
- Aligning one's decisions on socially accepted behavior

Examples of cognitive actions

- **Controlled memory** (versus automatic memory)
- **Perceptual discrimination**
- Problem **solving**
- **Explaining-why**
- Deliberating
- Questioning

Scientific metacognition has the potential to impact central philosophical issues

1. Naturalize and redefine "mental actions"
2. Naturalize epistemic norms of evaluation and develop cognitive epistemology
3. Document self-awareness in nonhumans

Outline

1. Epistemic agency

- Definition
- Informational goals
- Homeostatic structure

2. Epistemic or instrumental norms? Relevance of metacognition to epistemology

3. Variety of epistemic norms: from biology to cultural evolution?

1. Epistemic agency

How to define a cognitive action

Defining action

An action consists in **using compensatory mechanisms to maintain a trajectory toward a goal**

*From Harry Frankfurt "The problem of action"
(1978)*



The novelty of Frankfurt's definition

- Prior intentions are no longer a **defining** feature of action.
- This view is consistent with the fact that agents select informational goals
 - without having formed a conscious intention to engage in them (see verbal communication, error correction, etc.)
 - Without representing their informational goal in words (infants, nonhuman animals)

Defining a **cognitive** action

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

From Harry Frankfurt "The problem of action" (1978)



What is an **informational** goal?

Cognizers need to

- **acquire** information from their environment
- **retrieve** information from memory
- **reason** on the basis of what they know

→ Selecting important goals and obtaining correct, unambiguous information as a function of the resources available is a basic function of all animal minds called "metacognition"

Examples of informational goals

- Remembering who attended a meeting
- Identifying a bird, an insect, a flower, a tree
- Finding out where food is
- Solving a math problem
- Explaining a natural or a social phenomenon

Framework for metacognition

Nelson and Narens 1990

- **CONTROL**

- selecting a cognitive action
- adjusting effort to expected gain,
- modifying commands as a function of obstacles encountered on the way to the goal
- stopping
- channelling outcomes to other mental areas.

- **MONITORING**

Feedback from performance is used to predict discrepancies between expected and observed cognitive outcomes **at critical junctures** of the task.

3 Critical junctures: predictions about

1. Goals

- information as available: **curiosity**
- Feasibility: **feeling of knowing, ease of processing**
- success : **feeling of confidence**

2. Progress toward the goal : **feelings of understanding/confusion, feelings of error**

3. Outcome : **postdictive feelings of confidence, feeling of being right.**

Taxonomy of metacognitive feelings based on activity segments being assessed (reproduced from Goupil & Proust, *Cognition*, 2023)

Goal-related predictive feelings	Process-related evaluative feelings	Result-related evaluative feelings
Feelings of curiosity	Feelings of error	Feeling of being right/wrong
Feelings of familiarity	Feelings of incomprehension	Feeling that one learned ("judgement of learning")
Feelings of knowing	Feelings of incoherence	Eureka feeling
Feelings of prospective confidence	Feelings of interest/boredom	Feelings of retrospective confidence
Tip-of-the tongue	Feelings of confusion	

Monitoring an epistemic goal depends on the goal pursued

- Example:
 - Goal is to remember **accurately** who was there at a meeting: correction requires no false positives, but tolerates omissions.
 - Goal is to remember **exhaustively** who was there at a meeting: correction tolerates false positives, but requires no omission.
- two different norms involved in evaluations at each action juncture

Monitoring generates feelings of (un)certainty

- Metaperception: confidence in valid discrimination
- Metamemory: confidence for retrieval
- Metareasoning: feeling of being right
- Metacomprehension, confidence in grasping the implications of a sentence.

Diverse phenomenology

- Scientists refer to "confidence" in connection to different tasks
- There seems to be transfer of feedback calibration across cognitive goals (Fleming 2023 for a review).
- However, feelings share valence , not the specific experience associated to them: compare **feeling of knowing, feeling of understanding** with **judgment of learning**

Feeling-based predictions

- are "normative" to the extent that they **objectively enhance the cognitive efficiency** of agents' decisions.
- They are **automatically calibrated** over time to maximise predictive accuracy by implicit learning (i.e. the brain extracts the feedback from prior instances of a specific cognitive task)
- However, human agents can acquire beliefs and form theories that can **either improve or impair** epistemic decision-making

As seen above, this is made possible by the superposition, in humans, of 2 metacognitive evaluative processes

- **Procedural metacognition (PM)** is based on nonconscious predictive heuristics and comparators and their affective outputs (metacognitive feelings).
- **Explicit metacognition (EM)** is based on episodic memories, beliefs and theories about the on-going task and own ability in this task.

Koriat & Levy-Sadot (1999).

Explicit metacognition

- Evaluative, concept-based judgments focusing on cooperative epistemic tasks and justification (Frith, 2012)
- serves epistemic ends, such as promoting sophisticated knowledge goals (e.g. evidentiality, concept dependability) (Shea, 2020)
- also serves social functions, for example structuring epistemic labor in special fields (e.g. legal/scientific reasoning).

2. Epistemic norms

From the psychology of metacognition to epistemology

Two main issues

- A. What is meant by "an epistemic norm" ?

- B. How do informational norms (eg. truth) relate to rational norms (utility) ?

A - What is meant by
"an epistemic **norm**"?

In what sense can we speak of "norms" in metacognition?

- As seen above, predictions are called "normative" *to the extent that they objectively enhance the cognitive efficiency of agents' decisions as a consequence of a reliable informational (= predictive) mechanism (Proust, 2013)*
- Predictive mechanisms include
 - contextually reliable cues (procedural metacognition)
 - theories , beliefs or social prescriptions (explicit metacognition)

The duality in **evaluation** systems raises the issue of domains of norm validity

- For example, **a feeling of knowing** tends to reliably predict the ability to **retrieve a knowledge** item ("What is the capital of Australia?")
- But it does not specifically predict the **accuracy** of one's response ("consensuality principle", **Koriat 2008**)
- Feelings are more reliable in some contexts than in others (**Undorf, Navarro- Báez & Zimdahl, 2023**).

Is **explicit metacognition** more reliable?

- Not when theories relied upon are wrong:
 - "Learning is more efficient when one studies during two hours than during four half-hours"
 - "Girls are generally poor performers in math"
 - "Success at school only depends on students' efforts"

Cognitive agents learn to use "normatively predictive" cues

- They may individually learn over time to discriminate reliable from illusory predictors
- **Social metacognition** : institutions (legal system, education, scientific research) may disseminate reliable methods for inquiry and reasoning and normatively shape epistemic goals

Supra-personal cognitive control and metacognition

Nicholas Shea¹, Annika Boldt², Dan Bang^{2,3,4}, Nick Yeung²,
Cecilia Heyes⁵, and Chris D. Frith^{4,6}

¹Department of Philosophy, King's College London, Strand, London WC2R 2LS, UK

- "Humans evolved the ability to explicitly represent properties of their cognitive processes: by sharing and discussing these representations, agents can engage in novel forms of adaptive group behaviour and build cumulative culture."

Trends in cognitive sciences, 2014, 18(4)

Supra-personal control allows **socially shared** epistemic norms to emerge

Education, legal systems, and science organisations have significant epistemic roles (Goldman, 1999)

They may also be epistemically harmful by

- Unduly reducing the scope of epistemic goals (remembering words//understanding situations)
- Ignoring the affective dimension of learning

Supra-personal control allows **socially extended** epistemic norms to emerge

Socially accepted normative practices include **extended components, such as internet resources and outsourced constituents.**

People are **confident to know P**

- because they read P on internet
- because they know how to find internet confirmation relative to P.

Supra-personal control allows **socially extended** epistemic norms to emerge

The sensitivity to justification (evidence-based/inference-based) for having knowledge is a matter of dominant cognitive practices in a group.

In many communication contexts (Facebook, Twitter), **fluency** (of contents) and **metafluency** (of cognitive habits) are deemed to provide a form of justification ("post-truth") (Reber & Norenzayan, 2018)

Ambiguity of fluency as a predictor

- What is **quickly processed** generates a pleasant feeling of successful processing.
- Performing **the same type of cognitive action** as others in a context reinforces confidence in its appropriateness
 - Malagasy acceptances about mental activity after deceased family members vary from hospital to burial contexts ([Astuti, & Harris, 2008](#)).

Is fluency an **epistemic** norm?

- It may rather work as a resource constraint being functionally applied to any action.
- Hence its proper function is one of control, not monitoring.
- This leads us to our following research question

B - How to distinguish informational from instrumental norms?

A preliminary distinction: Detecting, believing, accepting

- **Detecting** involves extracting from perceptual input a predictive cue able to guide an adaptive behavior. (e.g. self-other produced movement)
- **Believing** consists in attributing a high probability for a proposition to be true and revising one's expectations accordingly
- **Accepting/judging** is the outcome of a cognitive decisional process

A traditional puzzle in epistemology

- Pragmatic philosophers consider **that acceptance of a proposition (the positive outcome of an evaluation)** is governed both by epistemic and instrumental norms (Stalnaker, 1984, Wray, 2001)
- This threatens the **autonomy** of epistemic evaluations (what if *truth* partly depends on utility?)

First observation

- An epistemic action is usually embedded in an instrumental (world-directed) action. For example:
- In order to shop for food, I need to remember the items on my shopping list (which I forgot to bring with me).

Epistemic actions are embedded in pragmatic (instrumental) actions

Epistemic action:
Epistemic norm(s)



Instrumental action: norm of utility

Consequence of embedding

- One's cognitive goal instrumentally depends on one's practical goal

→ The epistemic norm guiding a cognitive action is **selected** on the basis of the practical goal of the world-directed action

Second: From epistemic to strategic acceptances

Why strategic acceptance ?

- A rational subject may or not decide to act on his/her epistemic acceptance, **depending on the risk and benefit at stake.**
 - Utility does not just influence the **selection** of certain epistemic norms of acceptance.
 - It also influences **decisions to act**, when openly expressing one's own cognitive decisions would involve high risks or jeopardizes instrumental goals.

Proust, 2013

Epistemic action:
Epistemic norm(s)



Instrumental action: norm of utility

Argument from metacognitive studies (Koriat and Goldsmith, 1996)

- In situations where agents **are forced** to perform a cognitive task, strategic acceptance is ruled out: agents merely express their epistemic acceptance.
- In contrast, when agents **can freely consider how to plan their action**, given its stakes, they can refrain from **acting** (incl. **reporting**) on the unique basis of their epistemic acceptance.

Argument from metacognitive studies

Agents are allowed **to strategically withhold or volunteer an answer** according to their personal control policy (risk-averse or risk-seeking), associated with the anticipated costs and benefits
(Koriat and Goldsmith, 1996).

Argument from cognitive neuropsychology

- Strategic acceptance can be impaired in patients with schizophrenia, while epistemic acceptance is not **(Koren et al. 2006)**
- this confirms that epistemic and strategic acceptances are cognitively distinct steps.

In summary

Acceptances are context-dependent for two reasons:

1. **their norm** (constituting **this** type of accepting) is strategically dependent on an instrumental goal
2. **The decision to further act on epistemic acceptances** are strategically adjusted to the expected gain/cost of doing so.

The **Effortfulness** objection

- An epistemic action has its own cost and its own stakes: **is not this an indication that it needs to involve a dimension of utility?**
- Predicted reward influences effortfulness, i.e. the selection of a cognitive action and the resources allocated to it, including duration. (Kurzban et al., 2016, Lin et al., 2022)
- Hence, effort **controls the informational quality** of a cognitive action.

Utility norms contribute to modulating effort

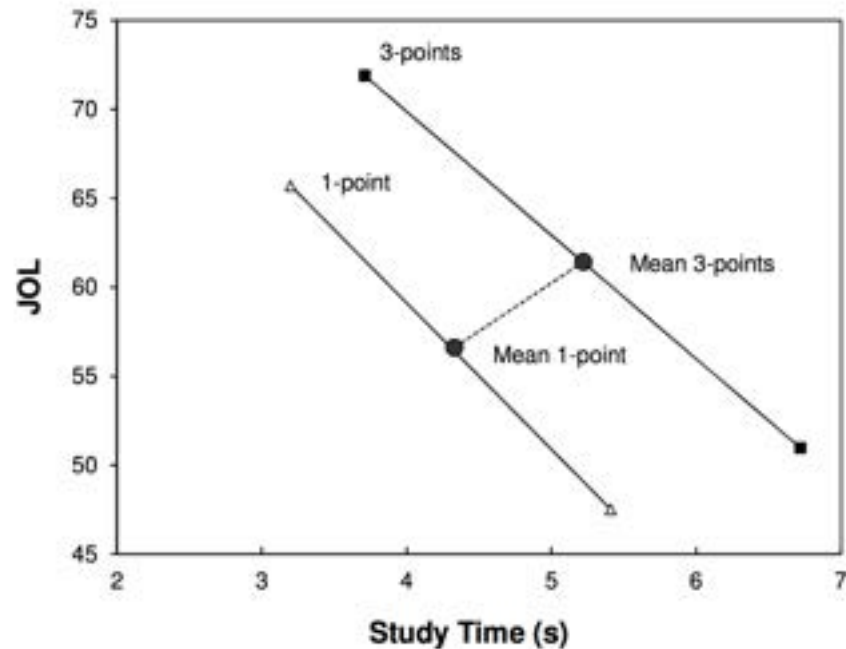
- Effort does correlate with higher reliability and higher confidence in one's current success (Koriat et al., 2006).
- This is so because **more resources** are invested in the task.

still, e-norms and u-norms are
separate evaluators

- Consider, for example, the judgments of learning in a learning task **that can be rewarded 1 or 3 points.**

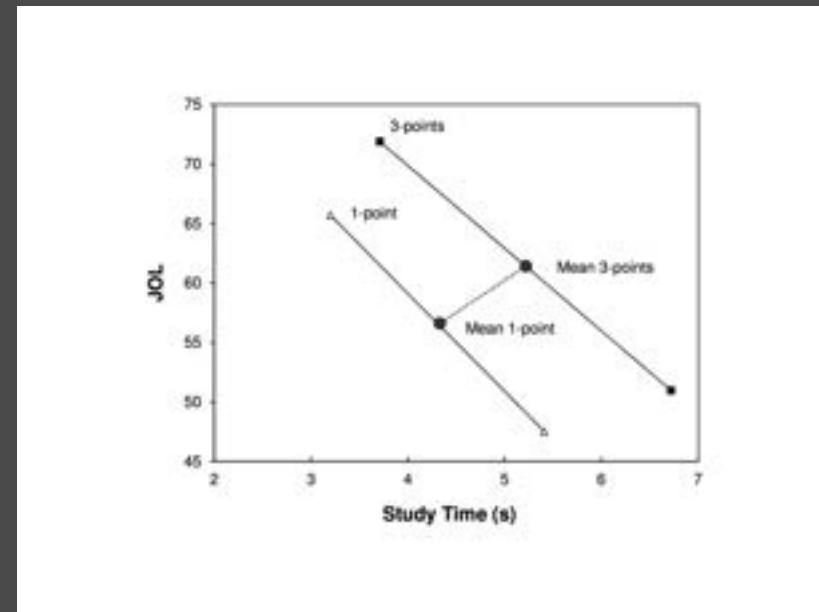
Koriat, A., Ma'ayan, H., & Nussinson, R. (2006).

Mean judgment of learning over 2 gain conditions (Koriat et al., 2006)



Evaluation process is independent from expected reward

- Participants spend more time learning, and have higher JOLs when they expect a 3 point reward than when they expect 1 point
- Their confidence, however, inversely covaries with study time, i.e. the actual cue on which learning predictions depend.
- The relative difficulty of each trial, then, is evaluated in the same way.



3. Variety in epistemic norms: from biology to cultural evolution

A wide variety of epistemic norms are used in the control of human cognitive activity

Epistemic Norms regulate types of acceptance (Proust, 2013)

- Fluency, beauty, precision (perception)
- Accuracy (memory, reasoning)
- Comprehensiveness or exhaustiveness (memory, reasoning, planning, problem solving)
- Consistency (fiction, deductive reasoning)
- Plausibility, likelihood (reasoning, planning)
- Intelligibility, explanatory power (understanding)
- Consensuality (negotiation)
- Relevance (problem solving, reasoning, communication)

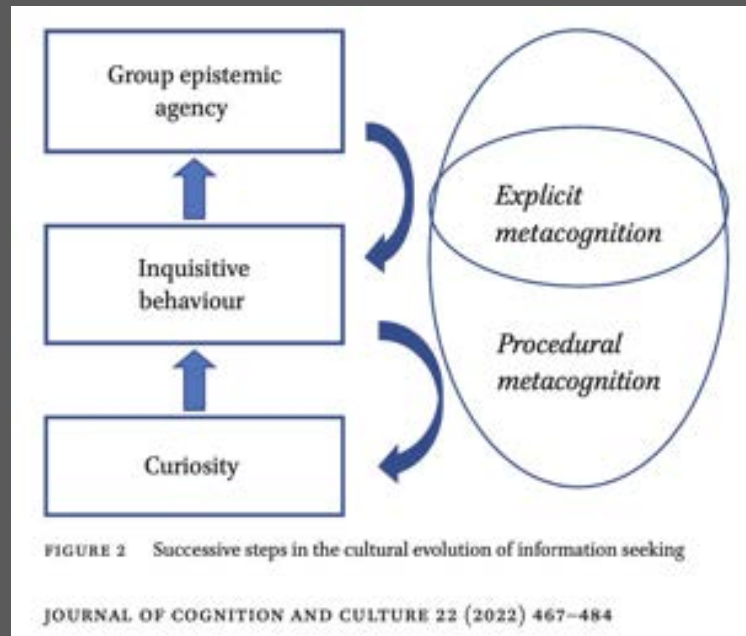
Nonhumans and young children

Monitor and control

- What they know about their environment
 - Their perceptual abilities (discrimination)
 - Their learning errors
 - Retrieval from memory
- Corresponding norm
 - **Informativeness** [infants] (Goupil & Kouider 2019)
 - **Perceptual accuracy** [rhesus monkeys] (Smith et al, 2003)
 - **Via Fluency** [zebra finches] (Gadagkar, 2019)
 - **Accuracy via Self-consistency** (Koriat, 2008)

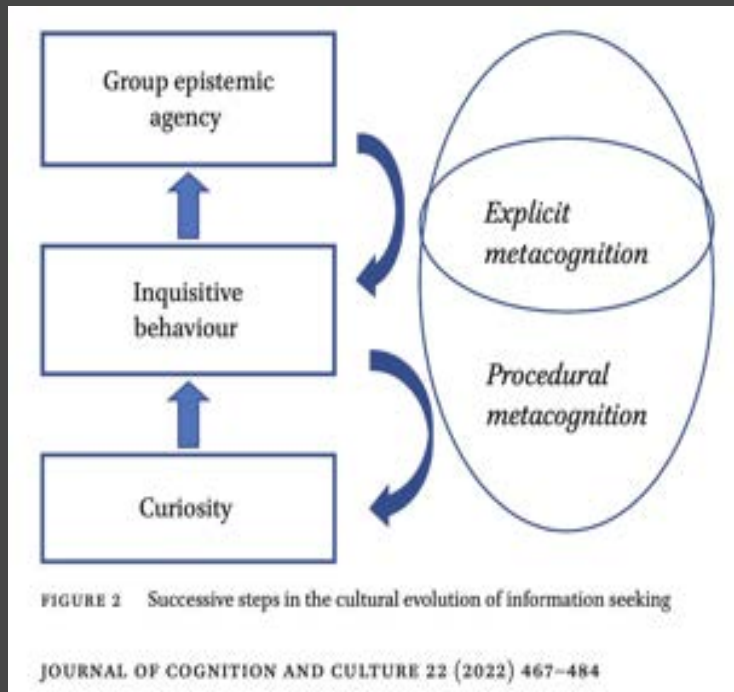
How did human sensitivity
to metacognitive norms
culturally evolve?

Hypothesis on the main phrases of the cultural accumulation of metacognition for information seeking (Work in progress)



Proust, 2022

How cultural accumulation might have developed

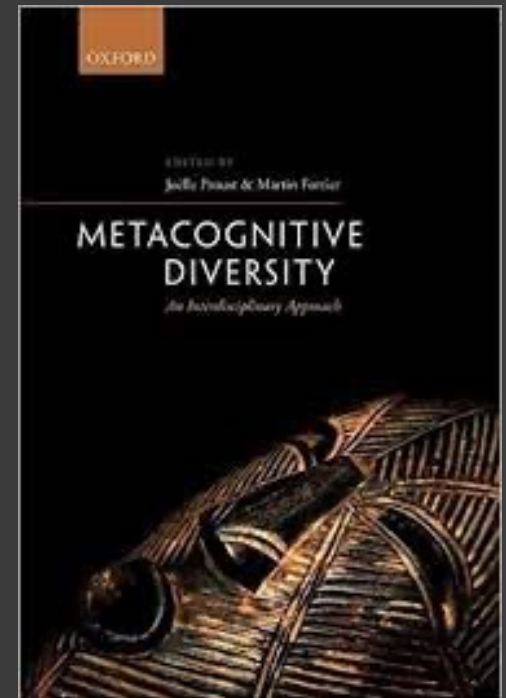


- Consensuality
- Confidence in plausibility
- Confidence in truth
- Conceptual understanding
- Relevance/explanatory power
- Feelings of perceptual or memorial confidence
 - Call discriminability
 - Feelings of knowing
 - Feelings of familiarity

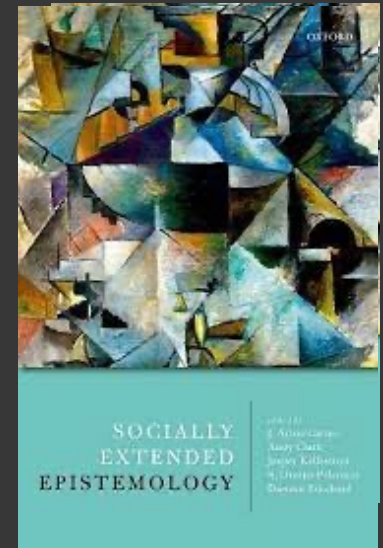
Feelings of fluency

Relevance of cognitive institutions to norm sensitivity

- Shamanism
- Oracles
- Religious credos and textual studies
- Public education systems
- Legal systems and courts
- Learned societies
- Scientific journals



From mindreading to supra-personally extended epistemology



Meta-representing one's own and others' beliefs and intentions helped

- differentiate and transmit cognitive goals
- broadcast cognitive outcomes → teach, learn more efficiently
- compare cognitive strategies
- detect potential trustworthiness, expertise and genuine epistemic commitment.
- New technologies, however, may allow collective decision making to facilitate access to relevant knowledge (with well-known potential costs). see Carter et al. *Socially extended epistemology*, 2018)

Procedural metacognition remains at the heart of human thought

- Explicit metacognition needs **affective evaluations** to guide the exploration and communication of one's own decisions
- Uncertainty, curiosity, the feeling of progressing in learning, will remain essential drivers of individual thinking and collective action.