Curiosity as a metacognitive feeling
Louise Goupil, Joëlle Proust

To cite this version:
10.1016/j.cognition.2022.105325 . hal-03648661

HAL Id: hal-03648661
https://hal.science/hal-03648661
Submitted on 21 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Curiosity as a metacognitive feeling

Louise Goupil & Joëlle Proust

Author affiliation

Louise Goupil, LPNC, Université Grenoble Alpes, Grenoble, France.

https://www.louisegoupil.co.uk/ - louise.goupil@univ-grenoble-alpes.fr


https://joelleproust.org/en/ - joelle.proust@gmail.com

Keywords

curiosity, explorations, learning, metacognition, questioning.

Word counts

abstract: 228

main text: 11499

references: 3871

entire text: 15798
Abstract (228 words)

Curious information-seeking is known to be a key driver of learning, but characterizing this important psychological phenomenon remains a challenge. In this article, we argue that this requires qualifying the relationships between metacognition and curiosity. The idea that curiosity is a metacognitive competence has been resisted: many researchers have assumed both that young children and non-human animals can be genuinely curious, and that metacognition requires conceptual and culturally situated resources that are unavailable to young children and non-human animals. We suggest that this resistance is unwarranted given accumulating evidence that metacognition can be deployed procedurally, and defend the view that curiosity is a metacognitive feeling. Our metacognitive view singles out two monitoring steps as a triggering condition for curiosity: evaluating one’s own informational needs, and predicting the likelihood that explorations of the proximate environment afford sizeable information gains. We review empirical evidence and computational models of curiosity, and show that they fit well with this metacognitive account, while on the contrary, they remain difficult to explain by a competing account according to which curiosity is a basic attitude of questioning. Finally, we propose a new way to construe the relationships between curiosity and the human-specific communicative practice of questioning, discuss the issue of how children may learn to express their curiosity through interactions with others,
and conclude by briefly exploring the implications of our proposal for educational practices.
1) Introduction

Curious information-seeking has long been identified as a key driver of learning (Baldwin & Moses, 1996; Berlyne, 1978; Bruner, 1961; Dember & Earl, 1957), and exciting empirical research has largely confirmed this assumption by gathering data in human adults, children and infants, as well as in non-human primates (Begus & Southgate, 2018; Gottlieb & Oudeyer, 2018; Gruber, Gelman, & Ranganath, 2014; Kang et al., 2009; Kidd & Hayden, 2015). The status of curiosity, however, remains hotly debated. There is at present no consensus about how to define and operationalize this important psychological phenomenon, nor about its emergence across ontogeny and phylogeny (Begus & Southgate, 2018; Carruthers, 2018; Kidd & Hayden, 2015).

This article proposes that defining curiosity requires clarifying its relationship with metacognition, broadly defined as the ability to evaluate the quality of one’s own informational states, and the efficiency of one’s own learning attempts, in order to regulate subsequent cognitive activities and behavior. Many researchers have resisted the idea that curiosity is fundamentally a metacognitive competence, because they have assumed both that non-human animals and
young children can be genuinely curious - since they engage in information seeking even when this does not lead to immediately consumable rewards (Bazhydai & Westermann, 2021; Gottlieb & Oudeyer, 2018; Kang et al., 2009; Kidd & Hayden, 2015) - and that metacognition requires conceptual and culturally situated resources that are unavailable to young children and non-human animals (Bazhydai & Harris, 2021; Begus & Southgate, 2018; Carruthers, 2018a,b; Perner, 2012).

Here, we argue that this resistance is not warranted: a large body of research suggests instead that metacognition does not reduce to its conceptual, meta-representational forms, and that preverbal children and some nonhuman animals possess rudimentary metacognitive resources (Goupil & Kouider, 2019; Proust, 2012, 2019; Shea et al., 2014). In our view, intrinsic curiosity must involve minimal forms of metacognitive monitoring and regulation allowing agents to identify and satisfy their informational needs, which we refer to as "core" or "procedural" metacognition. Our suggestion is thus that, from early on in development, and across animal taxa, engaging in curious information seeking is, fundamentally, a metacognitive competence.

First, we will present a basic definition of curiosity and a few important functional distinctions. We will then introduce two competing theoretical accounts that both attempt to characterize the semantic structure of curiosity,
and its link with metacognition. According to a first account, curiosity is a basic affective attitude of questioning (BQA account, Carruthers, 2018a,b). According to our proposal, curiosity is a metacognitive feeling whose function is to motivate cognitive agents to adaptively fill their informational gaps (MF account). We then discuss relevant empirical and theoretical arguments, and conclude that a metacognitive account is to be preferred, both on theoretical and empirical grounds. Finally, we discuss the relationship between questioning and curiosity in cross-cultural contexts and over development, and briefly explore the implications of our proposal for educational practices.

2) What is curiosity?

2.a.) Basic definition and functional distinctions

Curiosity is the mental property that is pretheoretically used to explain why people actively seek information. Why does Ann ask a question about a given event? Because she is curious: she wants to ascertain whether an event happened, and/or to learn more about where, when and why it happened. Curiosity then, is a mental state whose function is to trigger information search, and that is extinguished when information is obtained (Berlyne, 1962; Kidd & Hayden, 2015; Loewenstein, 1994).
Curiosity leads to exploring (Berlyne, 1966), but it is worth noting that not all explorations are driven by curiosity. Explorations are behaviors whereby agents search their environment. They are ubiquitous across the animal kingdom, and can be driven by different motives. Utility-driven (or instrumental) explorations have a non-cognitive goal, such as foraging for food or shelter. Curiosity-driven explorations, by contrast, occur when agents seek information for its own sake. In experimental contexts, curiosity is hypothesized to lead to exploratory behaviors when animals are satiated and safe (Kidd & Hayden, 2015; Oudeyer & Smith, 2016).

This distinction between utility-driven and curiosity-driven explorations is blurred whenever instrumental explorations incidentally lead to information gains. Imagine that you just moved in a foreign country. In the morning, you try to find a coffee shop. During this utility-driven exploration, you may notice events that surprise you (i.e., that differ from the expectations you formed on the basis of your prior knowledge). For example, a surprising traditional outfit could pique your curiosity, and lead you to want to know more about it (e.g. its name, its social meaning). Here, your desire for coffee incidentally triggered unrelated curiosity-driven information search.
The two types of exploration have a distinctive selective history and hence, different primary functions (Millikan, 1989). It is likely that utility-based exploration was selected first, to allow organisms to decide whether to exploit identified resources, or to sacrifice current gains with the prospect of better future rewards. A further adaptation then allowed organisms to explore their environment to gain information in the absence of a current instrumental need, leading to curiosity-driven explorations. Their primary function is to enrich the model used for predicting future events, in order to reduce subjective uncertainty (Koechlin, 2014). In this view, curiosity is part of the complex and multifaceted set of abilities involved in computing the trade-off between exploring and exploiting (Schwartenbeck et al., 2019).

Our example above highlights surprise as a precursor to curiosity (Vogl, Pekrun, Murayama, & Loderer, 2019), and we discuss this possibility further in Appendix 1. Surprise is triggered when an unexpected event occurs, that is, an event that contradicts the agent's prior beliefs. The resulting prediction error can then enhance the agents' attention, leading them to revise their priors in an effort to minimize future prediction errors (Hohwy, 2012; Sim & Xu, 2019). Surprise and curiosity are often equated with one another (Kidd & Hayden, 2015; Schwartenbeck et al., 2019), but they are functionally distinct: surprise can lead to curiosity – as in our example above (see also Ligneul, Mermillod, & Morisseau, 2018) – but it can also trigger sequential utility-driven explorations:
if an animal is surprised not to find food in a predicted location A, it may search
for it in other locations out of sheer utility, rather than curiosity.

In summary, surprise-based sequential explorations stem from experiencing
local prediction errors, i.e., a divergence between internal predictions and
observed outcomes (Egner, Monti, & Summerfield, 2010). Their primary
function is to allow agents to notice unpredicted changes, and to locally adjust to
them. Curiosity-based explorations stem from detecting an informational need,
and are terminated once this informational need is satiated. They can be
triggered by surprise, when surprise leads agents to recognize that they have an
informational need, but can also be triggered by other mechanisms. Curiosity
primarily depends on an organism’s ability to become sensitive to its
informational needs, and the aim of a theoretical account of curiosity should
thus be to try and characterize how these informational needs are detected. This,
in our view, requires clarifying the links between curiosity and metacognition.

2.b.) Curiosity and metacognition: two views.

Two main views have been proposed to qualify the relationship between
curiosity and metacognition. A first view is that curiosity is a basic,
prelinguistic affective attitude of questioning, that is triggered automatically
when a cognitive activity fails to generate the expected outcome. A second view
is that curiosity is a metacognitive feeling that stems from evaluating one’s own
informational needs, and the likelihood that exploring will lead to information gains in a given context.

2.b.i Curiosity as a basic questioning attitude (BQA).

This view proposes that curiosity is akin to the process of questioning, and is one of its evolutionary and developmental preconditions. Building on a proposal by epistemologist Dennis Whitcomb (Whitcomb, 2010), Peter Carruthers (2018) observes that the content of a curious attitude (e.g., the question “where are my keys?”) specifies the set of propositional attitudes that would count as an answer (e.g., “the keys are in the drawer/in my bag/on the table”). Extinguishing curiosity requires selecting a correct answer in this set of alternatives, leading to the formation of a novel belief (e.g., “I believe the keys are in the drawer”). In this view, curious states are desire-like attitudes, because they have strong motivational and affective dimensions, but they differ from first-order desires because the agent seeks knowledge rather than consumption.

A key aspect of this view is that curiosity does not presuppose metarepresenting one's own knowledge, or lack thereof. For instance, the questioning attitude “where are the keys?” requires that the agent has a concept of “keys”, but not necessarily a concept of “knowing”. As such, basic questioning attitudes (BQAs thereafter) can be available to nonhuman animals and pre-linguistic children.
The defining feature of this affective state is that it has an interrogative content (expressed in words: “where is X?”): it is a kind of attitude. In contrast with metarepresentations, however, its content does not consist in embedded propositions such as "X wants to know whether it rains". Accordingly, it only requires mastering the first-order concepts that apply to the emotion-triggering situation (related, for example to predation, foraging, mating). Thus, note that in this view, curiosity does not require metacognition — equated by the author with the ability to form metarepresentations (Carruthers, 2016; Perner, 2012).

In summary, according to this account, only first-order concepts are needed to generate basic questioning attitudes, and to grasp and memorize answers to it. These affective attitudes of questioning are proposed to arise automatically when a cognitive activity (e.g., remembering) fails to generate the expected outcome (e.g., recovering a piece of information from memory).

2bii. Curiosity as a metacognitive feeling (MF)

On an alternative view, curiosity is a specific kind of metacognitive feeling that arises when agents evaluate that they have an informational need in a specific context. On this view, curiosity depends on two types of assessment, bearing respectively on the agents' own lack of knowledge and on the potential learning gains afforded by their proximate environment.
This proposal is informed by dual-process accounts of metacognition (see the Appendix 2 for details, and a defense of these accounts against the metarepresentational views mentioned above). Although they differ in their details, dual-process accounts propose that metacognition does not boil down to the ability to form metarepresentations, and does not need to rely on concepts - such as “knowing” or ”being certain” (Koriat & Levy-Sadot, 1999; Proust, 2012; Shea et al., 2014). Instead, they propose that metacognition encompasses all of the mental processes whereby agents internally evaluate the reliability of their cognitive activities (deciding, remembering, learning…), and use such information (e.g., a confidence, fluency or error signals) to regulate subsequent cognitive activities. These evaluative processes give rise to metacognitive feelings, which in turn initiate exploratory or corrective epistemic behaviors. On this view, metacognition can thus be deployed non-conceptually – in the form of “subjective experiences with a distinct embodied phenomenal quality and a formal object” of varying valence and intensity (Proust, 2015) - in order to evaluate informational needs or achievements. On the other hand, concept-based processes broaden the scope of metacognitive control by integrating considerations about tasks, individual competences, social environment etc. (A. Koriat & Levy-Sadot, 1999).

This article proposes that curiosity is a special type of metacognitive feeling. Many affective responses, such as the experience of pleasure in relaxing, or of
painful exertion, primarily stem from monitoring bodily states. Metacognitive feelings, however, are affective responses that stem from monitoring the success or failure of one's own cognitive actions (identifying, discriminating, recognizing, remembering, learning, etc.). Their generic function is to predict feasibility or success in various types of cognitive tasks (in perception, memory, reasoning…) based on different temporal segments of the cognitive activity (predictive, current, retrospective). As shown in Table 1, each type of metacognitive feeling motivates a specific kind of control: launching a cognitive task, revising it, accepting or rejecting its outcome. Curiosity belongs to the predictive kind of metacognitive feelings: one that is concerned with identifying informational needs and strategically planning explorations to satiate those needs.

<table>
<thead>
<tr>
<th>Goal-related predictive feelings</th>
<th>Process-related evaluative feelings</th>
<th>Result-related evaluative feelings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feelings of curiosity</td>
<td>Feelings of error</td>
<td>Feeling of being right/wrong</td>
</tr>
<tr>
<td>Feelings of familiarity</td>
<td>Feelings of incomprehension</td>
<td>Feeling that one learned (&quot;judgement of learning&quot;)</td>
</tr>
<tr>
<td>Feelings of knowing</td>
<td>Feelings of incoherence</td>
<td>Eureka feeling</td>
</tr>
<tr>
<td>Feelings of prospective confidence</td>
<td>Feelings of interest/boredom</td>
<td>Feelings of retrospective confidence</td>
</tr>
<tr>
<td>Tip-of-the tongue</td>
<td>Feelings of confusion</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Taxonomy of metacognitive feelings based on activity segments being assessed. This taxonomy manifests the similarity between feelings of curiosity, feelings of knowing (Koriat, 1995), feelings of familiarity (where a face, for example, motivates searching who that person is (Whittlesea, 2001) and tip-of-the tongue-phenomena (Frazier et al, 2021). They all reflect a tension...
that motivates further search; their positive valence expresses the promise of closing an informational gap.

261

2622.c. Divergences between the basic questioning attitudes and the metacognitive accounts.

264Do the two views really differ? At first sight, characterizing curiosity as a basic affective attitude might seem congruent with proposing that curiosity is a metacognitive feeling, and both BQAs and MFs are claimed to be present in nonhumans and in prelinguistic children. The two views differ, however, on two important issues. Their first divergence bears on the definition of metacognition. The second divergence relates to functional differences between basic questioning attitudes and metacognitive feelings.

271

2722.c.i) Divergence 1: Two ways of defining metacognition.

273The BQA account is part of wider explanatory framework in which metacognition is defined as cognition about one's cognition, and equated with the ability to form metarepresentations (i.e., propositional attitudes about propositional attitudes such as “I believe that [I remember P]”). As a consequence, only conceptual forms of self-attribute of knowledge or uncertainty are worth the name of metacognition (Carruthers, 2016, 2018; Perner, 2012, but see Carruthers, 2020). On the MF view informed by dual-processing accounts of metacognition, in contrast, feelings generated by the
evaluation of cognitive actions also qualify as metacognitive. A detailed defense of the dual-processing view of metacognition is provided in Appendix 2. We argue that the notion of metacognitive feeling is a necessary intermediary explanatory notion to guide research on curiosity, because it can explain how curiosity can be deployed non-conceptually while still possessing subject-specific and context-specific features. Hence, metacognitive feelings can explain the transition from first-order evaluations to concept-based metarepresentations. This is described in greater details in section 3).

2912.c.ii) Divergence 2: Affective attitude versus metacognitive feelings.

One might still argue that, aside from the definitional issue, BQA as presented in Carruthers (2018) has much in common with a view emphasizing the role of metacognitive feelings. Both accounts involve a prelinguistic, affective mental construct whose function is to drive intrinsic exploratory behavior in nonhumans, young and adult humans. There are three important functional differences between the two accounts, however, which lead to very different implications and predictions.

First, the two constructs are based on different underlying mechanisms. BQA draws on an analogy with semantic constructs such as beliefs and desires. Just as having a belief is the (non-metacognitive) causal precondition for making an assertion, having a desire-like attitude to know something is the (non-
metacognitive) causal precondition for deploying information-seeking. As an attitude, BQA needs to have a (first-order) concept-based propositional content: "Why is this person wearing this strange outfit"?

In contrast, dual-process theorists of curiosity do not endorse the view that curiosity needs to involve propositional contents. The basic ontology of MF rather includes predictions and evaluations – i.e., comparators, a calibration process, and the associative cues that can enhance or influence predictions. No restriction, then, is imposed on the representational input to curiosity. Curiosity can stem from a mere feeling of surprise triggered by an unexpected stimulus (Ligneul et al., 2018; Vogl et al., 2019), or by a feeling of familiarity (Whittlesea & Williams, 2001), a tip of the tongue phenomenon (Frazier, Schwartz, & Metcalfe, 2021), etc. Graded MFs primarily stem from consciously experiencing these evaluative processes and their outputs, rather than forming propositional attitudes involving first- and second-order concepts.

A second difference consists in the semantic relations between curiosity and its triggering condition. On the basic questioning view, being curious about the name of the capital of New Zealand is an affective attitude toward a desired piece of information, conceptually represented through first-order concepts (“the capital of New Zealand is…”). Yet, BQA does not say much about how a specific trigger generates an adaptive exploration. It only proposes that a BQA is
automatically generated when a cognitive action (e.g., a remembering attempt) fails. Instead of this all-or-none mechanism, MF proposes instead that curiosity, like any other (metacognitive) feeling, primarily derives from experiencing analog evaluative processes, and as such, has a *graded valence and intensity*. This graded nature allows organisms to select a specific cognitive focus, and to modulate the intensity of their exploring behaviors, as a function of their prior knowledge, and of the informational gain afforded by the specific context they are in.

A third difference relates to the issue of whether curiosity is a basic or a complex ability. BQA argues that curiosity is basic, which means that it is irreducible to further drives and semantic contents. By contrast, the metacognitive account proposes that curiosity stems from an open set of evaluative processes that assess subjective uncertainty, detect prediction errors, and motivate exploration as a function of its expected positive or negative value (Gottlieb, Oudeyer, Lopes, & Baranes, 2013; Litman, 2008). As we describe in greater details below, such combined sensitivity to context and prior knowledge is most clearly manifested in the U-shape of curiosity: curiosity requires a minimal amount of knowledge to arise, and it stops when the information need is satisfied (Kang et al., 2009).
In the preceding section, we introduced two main theories of the semantic structure of curiosity. They diverge on three crucial issues: 1) the constitutive role of first-order concepts versus evaluative processes; 2) the interrogative, all-or-none versus the evaluative, graded structure of curiosity; 3) the basic versus complex (i.e., subject- and context-sensitive) nature of curiosity. In this section, the properties of curiosity are analysed along these three axes, on the basis of developmental, experimental and computational evidence. Curiosity is shown to include two key monitoring steps: 1) an evaluation of the internally available information, and 2) an evaluation of the potential information gains afforded by a specific context. The evidence that curiosity is supported by graded, subject- and context-sensitive evaluative processes is compatible with the idea that curiosity is a metacognitive feeling rather than a basic questioning attitude.

**First step: monitoring prior states of knowledge**

Berlyne long ago hypothesized that specific forms of curiosity are meant to reduce the cognitive conflicts associated with subjective uncertainty (Berlyne, 1961). In a similar spirit, it was later proposed that curiosity is elicited by perceived informational gaps (Loewenstein, 1994). This theory predicts that agents will be curious when they encounter a situation in which several
alternative interpretations are compatible with the current situation, given agents’ prior beliefs. Agents then compare their confidence estimates for each alternative. Below a given confidence threshold, a drive to collect additional information arises (Golman & Loewenstein, 2018; Loewenstein, 1994).

This model is consistent with the key finding that self-reported curiosity depends on agents’ evaluations of their prior knowledge. “Feelings-of-knowing” and “tip-of-the-tongue” experiences (i.e., metacognitive feelings that occur when one is unable to recover an information from memory, but still feels that one knows it) are associated with greater attempts at recovering a memory, greater curiosity, and greater information seeking than “I don’t know” or “I know” states (Litman, 2005; Loewenstein, 1994; Maril, Simons, Weaver, & Schacter, 2005; Metcalfe, Schwartz, & Bloom, 2017). In situations where information gain is guaranteed - because exploring leads to encountering stimuli that are easy to process - human adults typically seek information when they have low confidence (Desender, Boldt, & Yeung, 2018). For instance, in a recent study, when an algorithm was trained to classify high versus low confidence responses from electroencephalographic data, it could also accurately predict whether participants would engage in information seeking or not (Desender, Murphy, Boldt, Verguts, & Yeung, 2019).

Low confidence already predicts information-seeking early on in development
For instance, in a perceptual task, 3- to 5-year-old children preferentially seek additional information – instead of responding by themselves – in conditions in which they also report low confidence on a picture-based scale (Coughlin et al., 2014). In the language domain, 4-year-olds’ curiosity about word meanings varies as a function of their confidence (Jimenez, 2018; Jimenez, Sun, & Saylor, 2018). In the memory domain, even 20-month-olds seek help when they forgot the location of a toy, to avoid recollection errors (Goupil, Romand-Monnier, & Kouider, 2016).

In summary, from early on in development, empirical evidence indicates that information seeking is supported by metacognitive monitoring processes that allow agents to evaluate their informational needs. Evaluating prior knowledge is thus a key factor driving curiosity, but it is only sufficient to guide curiosity in environments where explorations tend to lead to information gain. In an uncertain world, however, explorations can be fruitless. Hence, curiosity should also depend on the expected information gain afforded by a specific context for a specific learner.

Second step: trade-offs between prior knowledge and sources of knowledge.

It has long been known that even infants orient their attention as a function of how familiar a stimulus is to them (Fantz, 1964; Hunter, Ross, & Ames, 1982),
and more recent research has shown more specifically that infants tailor their attention as a function of how predictable and informative stimuli are (Kidd, Piantadosi, & Aslin, 2012; Poli, Serino, Mars, & Hunnius, 2020). Infants tend to look away from events that are either too easy to process (because they are highly predictable) or too difficult to process (because they are highly unpredictable). This attentional strategy characterizes infants' individual exploratory behavior toward visual as well as auditory stimuli (Kidd et al., 2012; Kidd, Piantadosi, & Aslin, 2014). Beyond stimulus predictability, a recent study found that the way infants pay attention to a visual display also depends on how much information can be gained from a specific observation, given all past observations of the display (Poli et al., 2020). Thus, from early on in development, attentional allocation depends on evaluations of potential information gains afforded by a specific context.

Combined with the monitoring step described above, these evaluative mechanisms can allow learners to adjust their explorations as a function both of the informational need they identified, and of the potential gain their proximate environment affords. As mentioned above, where exploration is guaranteed to result in information gains because it leads to discovering stimuli that are easy to process, low-confidence is linearly associated with information-seeking (Coughlin et al., 2014; Desender et al., 2018). By contrast, contextual variations in the likelihood that exploring will provide information gains leads to a U-
shape function: curiosity is maximal when participants have middle-range levels of confidence (Kang et al., 2009).

This U-shape function can be accounted for by complexity theory (Dember & Earl, 1957; Kidd & Hayden, 2015), which proposes that curiosity varies with a system's ability to absorb new information. The best learning target is neither overly simple (already encoded into memory), nor “too disparate from existing representations already encoded into memory” (Kidd et al., 2012). Complexity theory matches Lev Vygostky’s notion of a zone of proximal development. Instructional guidance, Vygostky observed, is most efficient when learning goals are neither already met, nor too distant from the student’s existing acquisitions (Metcalf, Schwartz, & Eich, 2020; Vygotsky, 1978).

Progress learning theory. It has been objected to complexity theory that the notion of an "intermediate challenge" remains to be operationalized (Oudeyer, Gottlieb, & Lopes, 2016). Furthermore, there are many intermediate complexity stimuli that do not trigger explorations, especially in changing environments (Schmidhuber, 2010). A dynamic version of complexity theory, progress learning theory (PLT), initially developed in developmental robotics, claims that the brain, as a predictive machine, "is intrinsically motivated to pursue activities in which predictions are improving" (Luciw, Kompella, Kazeronian, & Schmidhuber, 2013; Oudeyer et al., 2016). Curiosity, on this view, is an intrinsinc
motivation whose function is to identify "niches of learning progress". Specialized predictive heuristics influence the intrinsic motivation to engage or disengage from an activity (Luciw et al., 2013). On the one hand, monitoring one's prediction error rates offers a way to assess task difficulty, which motivates exploring midly unfamiliar topics (in agreement with complexity theory). On the other hand, monitoring one's progress rate of learning (i.e. the temporal derivative of performance) dictates when to engage in, persist in, or disengage from learning when exploring would likely not lead to any information gain (Ten, Kaushik, Oudeyer, & Gottlieb, 2021). These two components map onto the two monitoring steps singled out in the MF model.

A complementary Bayesian approach specifies how the exploration-exploitation trade-off varies as a function of the types of uncertainty encountered in a given context, and evaluations of prior knowledge (Friston et al., 2017; Schwartenbeck et al., 2019). When reliable information is available, exploration should target either events to-be-predicted ("hidden state exploration"), or the causal, parametric structure of the model itself ("model parameter exploration"). When it is not, they should proceed randomly.

Curiosity as a rational behavior: A rational account of curiosity has also been defended in order to distinguish the subjective and objective forms of informational uncertainty at play in curiosity. It proposes that the value of
knowledge depends on the causal structure of the environment, and that the function of curiosity is to rationally seek information that is maximally useful for the agent in a given context (Dubey & Griffiths, 2020). This can explain that people are more curious about novel facts in changing environments where past events are less likely to recur in the future. It can also explain why the relationship between uncertainty and curiosity varies as a function of how instrumental it is to gather information in a given situation: when a decision must be made, and accuracy is important, it is rational to seek information when uncertainty is maximal. By contrast, if information can be collected without direct instrumental pressure, it can be rational to orient to intermediate levels of uncertainty that afford sizeable informational gains at lower processing costs. On this view too, the feelings associated with novelty, learning rate, complexity and other factors might all influence curiosity. The function of curiosity, again, consists in tracking stimuli likely to maximize the overall value of knowledge acquisition in a given type of environment. Curiosity is thus the set of mechanisms that approximates the rational behavior of seeking stimuli allowing agents to maximize their ability to make appropriate responses in the future.

Although successfully tested through computational simulations, developmental robotics, and adult game decisions, these theories await further evidence from developmental studies. As reviewed above, we have evidence that information-seeking relates to confidence (Coughlin et al., 2014; Goupil et al., 2016), and
attention allocation reflects the potential information gain afforded by a specific context (Kidd & Hayden, 2015; Poli et al., 2020) from early childhood. But when do children start integrating both kinds of evaluations to engage in curious-based explorations? This remains unknown. We examine this specific question in further details in Appendix 1.

Valence and intensity.

Another key aspect of curiosity - compatible with the idea that it constitutes a metacognitive feeling - is that it varies in valence and intensity. As observed by Berlyne (1978) and Loewenstein (1994), experiencing a conflict or detecting an informational gap may generate an aversive feeling, which in turn motivates the exploratory behavior predicted to suppress the unpleasant current feeling. Yet, exploration is an approach behavior because it predicts a positive experience of interest and the resolution of one's own perceived ignorance. Consistent with this, neural systems typically involved in reward anticipation also show increased activity for information anticipation during curiosity or curiosity relief (Gruber et al., 2014; Kang et al., 2009; Lau, Ozono, Kuratomi, Komiya, & Murayama, 2020; Ligneul et al., 2018).

The hybrid, "interest-deprivation" model (Litman, 2009) thus hypothesizes that curiosity is a trade-off between opposed motivations: a positive motivation for learning – interest – and a negative emotion generated by uncertainty and the
The respective weights of these affects reflect variations in learning goals. The trade-off might include additional motivations. The anticipated valence of the information to be gained has also been claimed to bias learning goals, and hence, curiosity about them. People seem to be less curious to know about predictably bad news than they are about good news (Gottlieb et al., 2013, but see FitzGibbon, Lau, & Murayama, 2020). This motivational articulation is compatible with the proposal that metacognitive feelings integrate predictive cues when monitoring a cognitive action from goal selection to action completion (see Table 1 above).

In summary, the evidence discussed in this section shows that, in line with the MF view, metacognitive monitoring is a key triggering condition of curiosity. Empirical and computational work indeed shows that curiosity expresses a trade-off between various motivations, each endowed with their own valence and intensity. This recognition speaks in favor of taking curiosity to be a metacognitive feeling that reflect subjective and strategic evaluations of informational needs, rather than constituting a basic affective attitude that is triggered automatically in an all-or-none fashion.

4) Curiosity and questioning
This being said, granting that in the BQA view, curiosity is a basic form of questioning, the MF view needs to clarify the relations between curiosity and the specifically human questioning behavior. The two theoretical accounts make radically different proposal regarding this aspect. In the BQA view, questioning and curiosity are essentially one and the same attitude. By contrast, in the MF view defended here, curiosity is a metacognitive feeling that can occur in the absence of any questioning ability. What additional functional steps, then, are involved when children start expressing their curiosity through questions? Before describing the relevant empirical data, the notion of questioning needs to be carefully examined. We will first analyze questioning as a communicational practice involving an inquisitive action, verbal or non-verbal, from a sender S to a receiver R, and clarify its semantic structure. We will then discuss the possibility that early, non-verbal forms of questioning express curiosity, and examine empirical evidence related to the question of whether and how questioning is scaffolded by interactions with caregivers.

4.a.) The semantic structure of questioning.

Questioning is often produced to reduce S's subjective uncertainty, or to test R's knowledge about a given subject matter, but it can also be used to provoke specific non-epistemic responses from R. Typically, in English, the function of the speech act "where is the salt?" is not to know where the salt is, but to request
the salt: it expresses a desire rather than curiosity. In Quichuan as in most other
languages, questions can be used to emphasize doubt to serve rhetorical
purposes rather than request information (Nuckolls & Swanson, 2018). Thus, S
may question R because *she wants to know something* from R (curiosity-driven,
epistemic questions), because she wants to deny the relevance of a potential
objection (pragmatic goal-driven, rhetorical questions), or more simply, because
*she wants something* from R (utility-driven questions).

Furthermore, articulating epistemic questions (verbally or nonverbally) requires
mastering the semantic structure for interrogative contents (Roberts, 2012). A
question not only has a focus, such as a new object in the environment; it also
represents a topic: a range of options to be narrowed down concerning the focus
(for example, how to name this object, or how to play with it?). Adequately
answering a question, then, requires identifying its focus along with the set of
relevant topical alternatives – or "contrastive topics" - that the questioner has in
mind.

4.b.) *From curious explorations to verbal questioning.*

---

1) Of note, this heterogeneity of question use is already visible during early childhood (Chouinard, 2007; Harris, 2020).
Interestingly, these properties of questioning are already apparent in early forms of pointing. At the end of the first year, infants start producing pointing gestures that are either utility-driven or curiosity-driven (Begus & Southgate, 2012; Tomasello, Carpenter, & Liszkowski, 2007), and that have specific focus, as manifested by a deictic (i.e. focused on an object) gesture, as well as a topic. Imperative pointings communicate a focused desire to get access to an object: "I want that!". In the context of joint attention, declarative pointings communicate a focused desire to share with another person the perception of an object of interest: "look at that" (Lucca, 2020; Tomasello et al., 2007). Infants also use pointing to communicate a focused curiosity – for learning about an object. Just like preschool-aged children keep on asking questions until they receive an explanation (Frazier, Gelman, & Wellman, 2009) and adjust their questions to their informational goal (Ruggeri & Lombrozo, 2015), infants’ tendency to point towards a novel object depends on whether they usually receive information upon producing that gesture or not (Begus & Southgate, 2012; Kovács, Tauzin, Téglás, Gergely, & Csibra, 2014). In addition, infants learn better the function and names of objects they have previously pointed towards (Begus, Gliga, & Southgate, 2014; Lucca & Wilbourn, 2016), which is also consistent with the idea they use these gestures to request information. We also know that 18-month children expect to receive information about object names in response to their pointing, not about their function (Lucca & Wilbourn, 2019). Granting that, in all these cases, inquisitive pointing expresses both a focus and a topic of
uncertainty, these gestures appear to be the first developmental expression of a genuine questioning ability.\(^2\)

Does this depend on a capacity to metarepresent knowledge and ignorance? Or do toddlers’ questions initially rely on non-conceptual forms of metacognition (i.e., on a metacognitive feeling of curiosity)? It has been proposed that questioning always presupposes an ability to conceptually represent the contrast between present ignorance and a specific knowledge goal (Bromberger, 1988). MF, however (as also does BQA), takes questioning not to presuppose a conceptual understanding of knowledge. Questions only involve concept-based metarepresentations if they focus on mental states (“what do you think about this”? "Are you sure to know where the toy is?"). Two lines of evidence are consistent with this idea.

The first shows that young children answer a question more reliably if it is not framed as a verbal report about what they know (Kim et al., 2016). Experiments such as the classical appearance-reality task demonstrate that 3-year-olds fail to track their own belief changes conceptually\(^3\) (Gopnik & Astington, 1988). Such

\(^2\) Research on social referencing (Bazhydai, Westermann, & Parise, 2020; Goupil et al., 2016; Hembacher, deMayo, & Frank, 2020) and object directed babbling (Donnellan, Bannard, McGillion, Slocombe, & Matthews, 2020; Goldstein, Schwade, Briesch, & Syal, 2010) suggest other potential candidates, although more research is required to confirm whether these behaviors reflect both contrastive focus and topics, which would confirm that they reflect genuine questioning.

\(^3\) In a typical version of this task, having predicted that a Smarties box contained candies, participants find out that it actually contains pencils. When asked what they earlier thought was in the box, most children report that they had always known” that it contained pencils.
Concept-based metacognitive failures can also be observed in the partial knowledge paradigm (Rohwer, Kloo, & Perner, 2012). This paradigm collects answers to a knowledge question in three conditions: children either have had full visual access to one of two objects being hidden in a box, partial visual access (they saw two objects, one of which would be hidden out of sight), or no access at all. When asked in the partial knowledge condition whether they know what is in the box, young children typically fail to verbally report their ignorance. This type of response is intriguing, because non-verbal measures independently suggest that young children do experience metacognitive feelings of doubt (Balcomb & Gerken, 2008; Coughlin et al., 2014; Goupil et al., 2016).

A study inspired by a dual-process view of metacognition used the partial knowledge paradigm to explore a possible dissociation between verbal report and metacognitive feelings in this setting (Kim et al., 2016). Three and 4-year-old children were either asked to verbally report whether they knew which toy was in a box, or to decide to inform another person about it. In the partial knowledge condition, children were much more reliable in choosing to inform than in verbally reporting what they knew. This difference is consistent with a dual-process view of questioning: reliably answering a question does not depend on metarepresenting what one knows or does not know. It rather depends on experiencing metacognitive feelings. Metarepresenting oneself as knowing, on

---

11 No difference between reporting knowledge and decision to inform was found in the full and absent knowledge condition, in line with previous findings that even 3-year-olds are able to verbally report their ignorance in the absent knowledge condition (Rohwer, Kloo, & Perner, 2012).
this view, depends on concept-based metacognition, whose dependence on social norms and values is well-established (Frazier, Schwartz, & Metcalfe, 2021; Proust & Fortier, 2018).

The second type of evidence comes from another study focusing on older children and adults' answers to ambiguous questions – where the referent of the question needs to be inferred from what the questioner allegedly knows or ignores about a situation given their perceptual access (Aguirre, Brun, Reboul, & Mascaro, 2022). It demonstrates that when interpreting epistemic questions, both adults and children readily consider what the questioner can or cannot see, but not necessarily what they know that they do not know. Thus, even adults do not necessarily interpret questions in a "Socratic ignorance" mode (assessing what the questioner knows that he does not know, Bromberger, 1988). Furthermore, children's ability to interpret questions that have an ambiguous focus was found to not correlate with their competence in passing second-order false-belief tasks (Aguirre et al., 2022).

Taken together, these findings suggest that – at a basic level - questioning does not require metarepresenting knowledge and ignorance. It merely depends on a prior detection of an informational gap and a learning affordance, associated with the subjective experience of curiosity, and on representing, however vaguely, a domain of alternative options; this requires metacognitive
monitoring, but it does not require a concept-based metarepresentation of one's own knowledge. Neither does the basic ability to respond to it (verbally or otherwise).

4.c. Conversational scaffolding of questioning.

The preceding section raises an important research issue: how do children transition from feeling curious, to expressing non-verbal questions, and finally putting questions into words? Much remains to be known about the various cognitive and social factors involved (Lucca, 2020), but most authors take questioning to be shaped by parental responsivity and glossing, and exposure to specific sociocultural communicational practices (Butler, Ronfard, & Corriveau, 2020; Chouinard, 2007; Salomo & Liszkowski, 2013). Three types of parental interventions might be particularly relevant.

First, scaffolding infants' attention to their subjective uncertainty about a given focus might enhance children's dispositions to be curious about their environment, and indicate to them how they may communicate this to others. As shown in Chouinard's monograph, the gestural expression of curiosity or doubt by an infant is often glossed by the parent as a specific question (Chouinard, 2007). Caregivers also tend to ask more wh- and yes/no questions in response to pointing compared to other behaviors (Wu & Gros-Louis, 2015). This suggests that parental glossing can train infants to ask questions verbally by
systematically mapping metacognitive feelings to specific inquisitive behaviors, words and linguistic expressions, supporting a form of cross-situational learning applied to mental states rather than external physical objects.

Second, for a given focus, caregiver's glossing builds up classes of contrastive topics (e.g., do you want to know how this is called? do you want to know how this works?). Affective intonation may be especially efficient in establishing relevant information in complement of syntactic and semantic cues (Roberts, 2012). Compare, for example, utterances of "Did John eat the cake?", i.e., [did he – rather than Mary – eat the cake?], with "Did John eat the cake?" i.e., [did John eat the cake rather than the banana?]).

Third, caregivers' ability to provide satisfactory answers can train toddlers' cognitive habit of questioning (Butler et al., 2020), and inform children about who – amongst the person that surround them – is knowledgeable (Harris, 2020). What is conveyed, beyond a specific interactive episode, is a conversational practice, whose specific goals are shaped by cultural constraints (see next section). Through questioning and evaluating the informative gain brought by the answers they receive, children can learn to satisfy their curiosity in a context-sensitive fashion: their natural tendency to be curious can be reinforced or discouraged. For instance, children ask more follow-up questions when they receive explanations in response to an initial question (Frazier et al., 2020).
and children whose parents provide more explanations are also more likely to come up with their own explanation when they find that an answer was unsatisfactory (Kurkul & Corriveau, 2018, also see below). Earlier in development, correlations have been found between caregiver responsivity and infant gesture use (Ger, Altinok, Liszkowski, & Küntay, 2018).

On top of supporting the transition from non-verbal to verbal questioning, conversational scaffolding may also have a deeper consequence. Non-verbal questioning behaviors could be an important way through which children transition from relying on metacognitive feelings to using metarepresentations. This is because parents’ responses to these communicative acts provide discrete conceptual inputs (e.g., mentalistic words) that children can systematically map onto analog metacognitive feelings. Linguistic input may thus support the process through which a non-verbal questioning ability fueled by analog metacognitive feelings is enriched by concept-based, metarepresentational abilities. This hypothesis remains to be tested however.

4.d.) Questioning as a communicational practice shaped by culture.

There is ample evidence that culture shapes metacognition (Kim, Proust, & Shahaeian, 2018) and questioning (Callanan, Solis, Castañeda, & Jipson, 2020; Gauvain & Munroe, 2020). Diversity begins with how situations are encoded
and reasoned about. Different attention patterns are transmitted by caregivers to children: they are predominantly holistic in Asia (relation-centered) and analytic (object-centered) in Western societies (Nisbett, Choi, Peng, & Norenzayan, 2001). Similarly, although curiosity is a universal metacognitive feeling, children should be curious about different things across culture, and be more or less willing to ask questions of a certain kind, depending on the communicational pragmatics of their social group. Although social norms regulate everywhere the public expression of curiosity, implicit or explicit rules concerning acceptable questioning foci differ across culture. In many rural traditional societies, such as Malagasy (Ochs Keenan, 1976), Pacific islands – including Samoa – (Ochs & Schieffelin, 1982/2001; Robbins & Rumsey, 2008), Mopan Mayas of central America (Danziger, 2010) or Mexican Mayas (Le Guen, 2018), others' whereabouts, beliefs or intentions are unsuitable conversational topics. In these social groups, children progressively learn to use indirect ways for obtaining information about what others think or do (Le Guen, 2018). In Western families, in contrast, infants’ vocalizations and toddlers' actions and utterances are considered as intentional and directly and verbally glossed as such (Ochs & Schieffelin, 1982/2001; Chouinard, 2007).

Similar socio-cognitive constraints shape acceptable ways of expressing and answering questions. Communicative practices centered on nonverbal behaviors such as gaze direction, gestures, subtle facial and postural changes, are favored
in many traditional cultures for asking or responding (such as the Mayan and Indian families studied in Rogoff et al., 1993). In contrast, middle-class US and Turkish children are more likely to use speech, as do their respective caregivers (Rogoff et al., 1993). Questioning and responding would also crucially vary with more general constrains related to which partner – the informant or the child – typically takes the responsibility for learning in a specific cultural setting (Rogoff, 2003).

In spite of the pragmatic differences discussed above, the overall frequency, and broad topical categories (e.g., activity, location…) of children's questioning does not appear to greatly vary with culture or socio-economic status (Callanan et al., 2020; Gauvain, Munroe, & Beebe, 2013). For instance, a study reported that 3-to 5-year-old Garifuna children from Belize, Newars children from Nepal, Logoli children from Kenya and Samoan children ask as many information-seeking questions as Western children from the US (Gauvain & Munroe, 2020; Gauvain et al., 2013). This quantitative approach, however, revealed a difference in the type of questions children asked. Children from these small-scale traditional societies produced fewer explanatory questions than American children (Gauvain et al., 2013). There is no evidence that such cross-cultural differences in explanatory questioning behavior correlate with advantages in reasoning or understanding facts about the world. Instead of interpreting deviations from Western norms as deficits (which poses obvious problems, see
such differences can be construed as reflections of wider diversities in cultural norms and practices. Accordingly, Gauvain and Munroe (2020) interpret this difference as a conjunction of two main factors. One is the cultural value attached to the relations between children and adults: respect, a social complement of age-based social distance, requires refraining from requesting explanations from adults (an implicit rule that all Logoli children followed several decades ago, when evidence was collected). In Western societies, adults are eager to prepare their children to formal school education. Parents’ tutoring is seen as the best way to promote children's learning (Ochs & Schieffelin, 1982/2001). Educated Western parents are prone to ask pedagogical questions with known answers (Yu, Bonawitz, & Shafto, 2019). In contrast, children from traditional societies are expected to cooperate with adults in carrying out community tasks, but are left more autonomy in epistemic matters than young Westerners. They readily learn by observing adults or older children what they need to know about the causal structure of both physical and social properties (Gaskins & Paradise, 2010).

The frequency of explanatory questions also differs among Western children as a function of their parents' socio-economical status (SES) (Kurkul & Corriveau, 2018; Tizard, Hughes, Carmichael, & Pinkerton, 1983). For instance, a recent study found that, even though the ratio of explanatory and factual questions remains constant across groups, middle-SES children ask twice the number of
explanatory questions than low-SES children. Middle-SES caregivers are also more likely to provide explanations, while low-SES families use circular responses more frequently, e.g., "because that's the way it is" (Kurkul & Corriveau, 2018), which may discourage further explanatory questioning.

While the type of questions used by parents may not impact children's reasoning when they are given exploratory autonomy, it may not be the case within Western societies, where formal instruction prevails over observational learning, and where children's learning opportunities largely depends on their educational environment, including caregivers' availability and skills, as well as institutional resources. A lack of practice in explanatory questioning might thus partially explain low-SES children's disadvantage at school (Jones, Swaboda, & Ruggeri, 2020), where teachers typically expect that children should master this communicative practice (Gauvain & Munroe, 2020). Relatedly, educational studies have demonstrated that the type of questioning used at school predicts school achievement. College students who seek teachers' help by asking explanatory questions have higher achievements than those who do not: asking for solutions – requesting expedient help – correlates instead with low achievement (Karabenick, 2004; A. M. Ryan, Patrick, & Shim, 2005). Much remains to be understood about the factors that may facilitate – or impair – questioning in the classroom, as a function of children’s socio-demographic and cultural origin, and with the aim to better foster their learning in educational
settings that may enforce communicational practices that may – or may not – match their familial practices.

In summary, although children across the world have the same disposition to be curious, the social norms governing inquisitive communication and adult-children interactions deeply influence how children ask questions, in agreement with the idea that questioning depends on learned cultural practices adapted to local demands regarding communication and teaching.

5) Conclusion

In this paper, we argued that curiosity is a metacognitive feeling whose function is to motivate cognitive agents to adaptively fill their informational gaps. In conformity with current theorizing, as well as empirical and computational work about curiosity, the metacognitive view defended in this article singles out two monitoring steps: detecting a knowledge gap and identifying the probability of reducing this gap in a given context. When agents have experienced that curiosity leads to intrinsic rewards – the pleasure of learning –, they tend to be more curious, hence, explore and learn more (Oudeyer et al., 2016). Unfavorable learning environments, however, would symmetrically reduce curiosity, exploration and learning. Early forms of curiosity and inquisitive behaviors, then, can be analyzed as a form of self-regulatory metacognition:
monitoring their subjective uncertainty leads infants to strategically question knowledgeable others, which turns out to increase their learning rate. Epistemic questions, focusing on object properties or on explanatory relations, allow learners to reduce the distance to their epistemic goals (i.e. satisfy their curiosity toward a given focus), within the limits of culturally prevalent socio-cognitive constraints.

Given the documented link between efficient information-seeking and academic achievement, gaining better insights into the mechanisms that support curiosity and questioning during early childhood is crucial (Selmeczy, Ghetti, Zheng, Porter, & Trzesniewski, 2021; von Stumm, Hell, & Chamorro-Premuzic, 2011, Yu et al., 2018). Distinguishing experience-based (metacognitive feelings) from concept-based (metarepresentational) forms of curiosity can lead to diversified pedagogical practices, which together can enhance learners’ motivation.

On the one hand, feelings of curiosity can be elicited and nurtured early on during spontaneous interactions with caregivers, for instance during play, through simple observation and reinforcement mechanisms. Adaptive questioning practices can similarly enhance childrens’ curiosity, and reinforce their motivation to explore. If questioning is not conducted in accordance with the notion of proximal development, however, it tends to fail to induce it (Frazier et al., 2021).
Formal tutoring and explicit training, on the other hand, mainly rely on concept-based, controlled forms of curiosity. Some educational practices can be incompatible with curiosity, for instance when they emphasize teacher's authority, extrinsic motivation (for grades, rather than for knowledge) and social competition (R. M. Ryan & Deci, 2000). Emphasis on teacher-guided instruction over autonomous exploration, and on conformity over questioning and creativity, has resulted in Western teenagers being less avid learners (Hattie, 2011; Kuhn, Modrek, & Sandoval, 2020). Conversely, explicitly metarepresenting what is known versus what is yet to be discovered has been found to enhance students' curiosity and their ability both to remember what they learned and to reason about it (Bjork, 2018). A question-and-answer method has been shown to be more effective than a text to achieve these goals (Iordanou, Kuhn, Matos, Shi, & Hemberger, 2019; Kuhn et al., 2020). Finally, a technique consisting in addressing questions to oneself, called "self-explanation", first used to train school children to understand what they read (McNamara & Magliano, 2009), has been successfully generalized to other educational fields (Rittle-Johnson, Loehr, & Durkin, 2017).

This diagnosis from educational studies is compatible with the experimental and theoretical work discussed in this paper, and emphasizes the importance of
understanding what curiosity is, how it relates to questioning, and how its early development influences life-long motivations to learn.

APPENDIX 1: When and how does curiosity develop?

Young children already seek information as a function of what they already know, and how confident they are in what they know. This is specifically manifest in their selection of objects to play with. They engage in exploratory play as a function of how well they understand a specific situation, and as a function of how much information gain they can expect from it (Gweon, Pelton, Konopka, & Schulz, 2014; Lapidow et al., 2021; Schulz & Bonawitz, 2007; Siegel, Magid, Pelz, Tenenbaum, & Schulz, 2021). For instance, when they observe confounding evidence about how to activate a pop-up toy, five-year-old children prefer to keep on playing with this “old” toy rather than playing with a totally novel toy (Schulz & Bonawitz, 2007). Children also explore a novel toy more thoroughly when a teacher provided partial information, as opposed to full information, about its function (Gweon et al., 2014). In the language domain, 3-8 year-old children preferentially seek information about novel words when they
are introduced to them in an ambiguous context that lead to referential ambiguity (e.g., in the presence of two potential referents) (Zettersten & Saffran, 2020).

But when does curiosity emerge during human development, and perhaps most importantly, how? In the main text, we allude to the proposal that surprise-based explorations constitute a precursor for the development of curiosity-based explorations (Perez & Feigenson, 2020; Vogl et al., 2019). If surprise is a precursor of curiosity, then when in development do children transition from surprise-based to curiosity-based explorations?

We know that unpredictability enhances infant’s attention (Kidd et al., 2012; Meyer, Schaik, Poli, & Hunnius, 2022; Poli et al., 2020) and learning (Stahl & Feigenson, 2015), and that attentional engagement during exploration predicts how well infants learn novel object properties (Begus, Southgate, & Gliga, 2015). When 11-month-old infants observe events that violate their expectations about object properties (e.g., a toy car that flies instead of obeying the laws of gravity), they spend more time looking at these objects, and learn their novel properties better (Stahl & Feigenson, 2015). Furthermore, when given the opportunity to explore these objects, infants tend to try and reproduce the surprising event, which can be seen as an attempt to test hypotheses about the surprising object.
A rich interpretation of this type of behavior is that unexpected events pique infants’ curiosity, leading them to subsequently engage in selective explorations to seek explanations. A leaner explanation that does not involve genuine curiosity, however, is equally possible: surprise may automatically enhance infants’ attention towards objects that have unexpected properties given their prior knowledge. Observing surprising events may lead infants to revise their belief about the object’s properties, and they may try to repeat the surprising event because they find it particularly rewarding and/or arousing. That is, maybe all infants do initially is generate expectations, revise their beliefs when expectations turn out to be false, and orient their attention selectively as a function of stimulus predictability. Thus, on the basis of this type of evidence, we cannot conclude on whether infants are curious per se (i.e., want to know whether the toy car flies or not), or whether they are merely surprised.

When exactly - and how - infants start engaging in curiosity-driven explorations during development is thus an important open question. One possibility is that surprise sets the stage for the development of curiosity (Vogl et al., 2019): infants may progressively notice that when they engage in utility- and surprise-
driven explorations, they incidentally gain new knowledge (as in the clothing example described above), and as a consequence, start exploring directly for this purpose. A recent study supports this idea, as it found that infants’ surprise responses (i.e., longer looks) to unexpected events at 11 and 17-month-old was correlated with curious traits measured via parental reports at age 3 (including questions such as “my child devotes considerable effort trying to figure out things that are confusing or unclear”) (Perez & Feigenson, 2021). These are correlational findings, but they are at least compatible with the idea that surprise is a developmental precursor to curiosity. Of note, a recent study found that, while for adults curiosity mainly reflects the informational gain they expect in a given context (a complex evaluation involving the two monitoring steps we outlined above), 5- to 9-year-old children's curiosity correlates not only with expected learning, but also with higher objective uncertainty and lower surprise (Liquin, Callaway, & Lombrozo, 2021). It may thus be that younger children's curiosity mainly depends on fluence-based elicitors, i.e., feelings such as surprise, while adults' curiosity takes better advantage of computing learning opportunities. Alternatively, reliably assessing expected learning in a given case may change across development, and rely on various predictors integrating knowledge and utility, exploration and exploitation.

APPENDIX 2. How should metacognition be defined?
This appendix addresses the definitional debate between a single-process and a dual-process view of metacognition, whose main tenets were presented in section 2 in the main text. Arguments for each view will be summarized and discussed as follows. First, the reasons advanced in favour of a monistic, concept-based definition of metacognition will be articulated as three main objections against a dual view. After a critical discussion of these objections, the arguments for defending a dual-process definition of metacognition will be presented.

**ARGUMENTS FOR A SINGLE-PROCESS VIEW OF METACOGNITION**

1. *The first-order objection.*

A main reason for endorsing an exclusively conceptual approach to metacognition is the semantic analysis of the term "metacognition", where "metacognition" means "cognition about cognition", which implies that metacognition is to be equated with an ability to hold metarepresentations of one's own attitudes (Carruthers, 2016, 2018a,b; Perner, 2012; but see Carruthers, 2021). Endorsing this definition leads to take subjective experiences of uncertainty as being merely "first-order" (Carruthers, 2016), because they are not explicitly *about* one's own mental states (Carruthers, 2016). Appropriately second-order metacognitive representations need to involve metarepresentations,
i.e. explicit representations of one's own mental states and properties.


Under the admission of nonconceptual, affective forms of control as metacognitive, one should consider as metacognitive the multiple forms of forward models that subpersonally control and monitor our cognitive activity, including those that make sensorimotor control or multisensory conflict resolution possible. Granting that subpersonal forms of information (probability distributions) are used as input to metacognitive predictions, the objection goes, should not the multiple forms of forward models that subpersonally control and monitor our cognitive and motor activity also qualify as metacognitive (Carruthers 2016, 2018a,b)? For example, why should not the comparators that regulate conflicts across sensory channels, such as the visual-auditory conflict involved in the McGurk effect, be categorized as metacognitive (Nagel, 2014)?

3. The irrelevance objection (Carruthers, 2016).

This objection states that metacognitive monitoring is not involved in nonhumans' and infants’ decisions to seek additional information instead of committing to a choice. Activities such as seeking information, trying – or declining – to remember, to discriminate, do not need to involve any form of
procedural evaluation. Non-human animals or human infants may instead decline a task or seek information simply because they lack a specific memory of the stimulus. Hence, decisions concerning memory depend on memory alone, not on meta-memory:

"For that one is disposed to act in one way if one has a memory, and in another if one does not is just what it is to have or lack a memory. This just describes the normal first-order causal role of memory in the cognitive and decision-making processes of creatures that possess memory-states." (Carruthers, 2016, p.74).

In addition, decisions to perform or to decline a memory task are argued to be sensitive to the likelihood of reward, not to the likelihood of cognitive success (i.e., epistemic uncertainty) (Carruthers, 2016, 2018a,b). A more economical account of infants' or nonhumans' opt-out or information seeking responses, so the argument goes, is a simple working memory decision based on reward probability that exclusively depends on the underlying first-order state.

II. RESPONSES TO THE OBJECTIONS.

Are feelings of confidence first-order?
Adressing this objection requires clarifying what one means by "uncertainty". On the one hand, uncertainty is an objective property of the sensory input. The variability of relevant stimuli in the external world, for example, is an objective source of uncertainty. The assessment through which the brain represents probability distributions of specific inputs is clearly first-order. Granting that there is noise in each informational channel, part of the objective uncertainty originates in the perceptual system itself (Pouget, Drugowitsch, & Kepecs, 2016). This kind of uncertainty, as an objective characteristic of the sensory input, does not yet qualify as a metacognitive signal.

On the other hand, uncertainty can also become a subjective property of the cognitive system: the brain compares the variance in the vehicle to stored values (concerning for example onset of neural activity, convergence or disparity of predictions); on this basis, it predicts (or estimates) the probability that a perceptual or memorial task will be (or has been) correctly performed. In the context of perceptual decision-making for instance, empirical evidence and computational frameworks based on evidence-accumulation principles suggest that decisions are made when perceptual evidence reaches a given threshold, and that associated confidence relates to the distance between the decision threshold and the maximum amount of evidence accumulated post-decisionally (Pleskac & Busemeyer, 2010) or dynamically around (Pereira et al., 2021) or up until
the time of the decision. The output of such comparators is a metacognitive feeling of confidence or doubt. Some researchers propose to call it "confidence" in contrast with "uncertainty" (Pouget et al., 2016). Its function is to predict the feasibility, or to monitor the desirable correction of a cognitive action, in order to select and motivate the congruent behavior (A Koriat, 2000).

It should be observed that sensitivity to one's own subjective uncertainty does not merely result in retrospective feelings of confidence or doubt. It applies to the variety of metacognitive feelings and predictive functions summarized in Table 1 above. In summary, the information used to appraise the likelihood of one's own success in a given cognitive task (remembering, learning, etc.) consists in sets of probability distributions concerning predicted outcomes. While these predictions are performed through nonconscious Bayesian computations, they generate a conscious feeling with a given valence, intensity, and motivational strength, such as the feeling of retrospective confidence attached to a cognitive decision.

Based on these considerations, one can conclude that the evaluative states involved in metacognition do not track "first-order" signals, i.e. objective
uncertainty, to the extent that such evaluations predict epistemic feasibility or correction from the agent's viewpoint in an occurrent context. They are based on heuristics formed by the agent across prior encounters with the task, and reflect a subjective state rather than an objective property of sensory inputs.

The ubiquity objection: are all predictive models metacognitive?

Addressing this objection requires focusing more closely on the function of metacognition. While there are nonconscious processes for revising one's own behavior – such as keeping one's balance (Klaus et al., 2020) or revising one's hand trajectory (Pélisson, Prablanc, Goodale, & Jeannerod, 1986) –, metacognition has the function of evaluating one's own informational states in order to control them. In other words, metacognition is specialized in the informational issues related to cognitive activities such as deciding, remembering or problem solving. To be sure: metacognitive evaluations do not predict the likely consequences of actions (e.g. grasping, catching…) but of cognitive actions (e.g., remembering, deciding). Hence metacognition does not adjudicate between sensory inputs, nor does it control and monitor sensorimotor activity. Its specific role consists in detecting and incorporating relevant knowledge into one's present or future decisions and cognitive actions.
But why is this function not involved in resolving conflicts between sensorimotor channels, might the objector insist? A plausible response is that sensorimotor control is the evolutionary basis of cognitive control. In both cases, the dorsomedial frontal cortex (DMFC) stores forward models in order to predict sensorimotor and cognitive success. Forward models automatically compare expected and observed dynamics of an action, in particular its predicted timing. The neural structures implementing these forward models seem to have been later recruited for the control of thought, i.e., metacognitive control and monitoring (Egger, Remington, Chang, & Jazayeri, 2019; Ritz, Frömer, & Shenhav, 2020). The forward models for cognitive control, however, do not use sensorimotor cues to automatically correct behavior; nor do they result in automatic compromises between sensory sources; they rather attempt to assess informational quality with an eye to its predicted value (rewards and risks incurred) with respect to commitments to specific alternatives, in ways that are subject-specific, and relate to the agents’ current goals and motivations.

3. The irrelevance objection: are metacognitive feelings necessary to explain children and non-human animal’s information seeking?

This objection fails to account for two central features of young children’s use of opt-out, or information-seeking options (the discussion of comparative
evidence is left aside here as it has been addressed elsewhere, e.g., see Beran, Proust, 2019).

First, adults and 3-yo children's subjective reports of confidence are related to opt-out and information seeking decisions (Coughlin et al., 2014; Desender et al., 2018). Assuming that functionally equivalent behaviours in slightly younger children should involve completely different mechanisms is not parsimonious.

Second, children not only opt-out / seek advice when stimuli are ambiguous or when the task is difficult (i.e., as a function of objective uncertainty). They do so as a function of the quality of their memory or decisions (Balcomb & Gerken, 2008; Geurten & Bastin, 2018; Goupil et al., 2016), a subjective form of uncertainty. When 20-month-old children are provided with the opportunity to ask their caregiver for advice instead of pointing themselves towards the location of a hidden toy, they indeed predominantly use this option as memorization delay increases (Goupil et al., 2016). But in addition, and crucially, for equivalent memorization delays (i.e., similar objective task complexity), toddlers provided with the possibility to ask their caregiver for help (test group) make less errors than toddlers who were not shown that their caregiver could help (control group). If lack of memory – or, in Carruthers' words, "a basic questioning attitude" – drove these behaviors - rather than strategic adjustments relying on metacognitive feelings - one should reasonably
expect that children in the control group would also avoid responding whenever they "have a lack of memory". But they don’t: children in the control group tend to respond by themselves, or occasionally to look away; furthermore, when they (very rarely) turn towards their caregivers spontaneously, this behavior is totally unrelated to their task performance and/or to difficulty. Thus, the failure of the cognitive activity of remembering does not automatically trigger help seeking in this age group.

III. ARGUMENTS FOR A DUAL DEFINITION OF METACOGNITION

Our responses to the objections given above are consistent with the general arguments offered in favor of dual process theories. According to this family of theories, an impulsive evaluative "System 1" generates a quick response, based on independent parallel frugal heuristics. A deliberately controlled "System 2" handles its evaluations serially, takes time to operate, consumes more working memory resources than System 1, and heavily relies on conceptual resources. While the representations in System 2 are characteristically conscious and volitional, System 1 representations are typically unconscious and influence behavior automatically (Shea et al, 2014, Reder, 1996, but see Koriat, 2000).

More specific arguments in favor of metacognitive duality range from comparative and developmental evidence, to behavioural and neural
dissociations in human cognitive decision-making.

1. Comparative evidence indicates that nonhumans (rodents, primates) are able to reliably evaluate their perception or their memory in a predictive or a retrospective way (Beran, 2019; Couchman, Beran, Coutinho, Boomer, & David Smith, 2013; Kepecs & Mainen, 2012; Proust, 2019).

2. A hierarchy among human brain systems reflects the evolution of control mechanisms towards enhanced exploratory flexibility (Koechlin, Summerfield, E, & C, 2007; Rouault & Koechlin, 2018).

3. In human ontogeny, children display behavioral markers suggesting that they experience feelings of curiosity and confidence before they can conceptually represent themselves as having these mental states (Bazhydai & Westermann, 2021; Goupil & Kouider, 2019).

4. In human adults, dissociations have been documented in metacognitive monitoring according to task demands, between feeling-based and concept-based incompatible evaluations (Koriat & Ackerman, 2010; Nussinson & Koriat, 2008). Neural signatures likewise distinguish deliberate, conscious error monitoring from automatic, unconscious error monitoring (Nieuwenhuis, Ridderinkhof, Blom, Band, & Kok, 2001; Yeung & Summerfield, 2012).

5. Clinical research also suggests dissociations between core metacognitive regulatory processes and the metacognitive sensitivity
expressed in subjective reports (Charles et al., 2017; Metcalfe, van Snellenberg, DeRosse, Balsam, & Malhotra, 2012; Nicholson, Williams, Lind, Grainger, & Carruthers, 2020). For instance, patients with schizophrenia show altered error-monitoring when performing decisions on the basis of supraliminal stimuli, but preserved error monitoring for subliminal stimuli (Charles et al., 2017). Children and adolescents with autism spectrum disorder show impairments in mindreading tasks (that tap into metarepresentational abilities), but not in post-decisional gambling tasks (that require them to evaluate the likelihood that their decision was correct, a core metacognitive ability) (Nicholson et al., 2020, but see Koren, Seidman, Goldsmith, & Harvey, 2006).

6. In collective cognitive decision-making, individual epistemic feelings need to be collected for optimal reliability in deliberation, but sharing one’s confidence with others requires complex interpersonal calibration processes that are constantly adjusted and re-negotiated through conversational alignment (Bang et al., 2017; Fusaroli et al., 2012).

Collectively, these lines of evidence show that there is an important functional distinction between system-1/core/procedural versus system-2/deliberate/conceptual forms of metacognition (Goupil & Kouider,
Acknowledgments. The authors thank Katarina Begus for helpful comments on a previous version of the manuscript.

Funding statement. This work was supported by the institutional grant ANR-17-EURE-0017 FrontCog.

Conflicts of interest: none.


Begus, K., & Southgate, V. (2018). Curious learners: How infants’ motivation to learn shapes and is shaped by infants’ interactions with the social world. In M. M. S. A. Ganea (Ed.), *Active Learning from Infancy to Childhood: Social Motivation, Cognition, and Linguistic Mechanisms* (pp. 13–37). Springer. https://doi.org/10.1007/978-3-319-77182-3_2


https://doi.org/10.1037/H0046861


https://doi.org/10.1523/JNEUROSCI.2620-18.2019


https://doi.org/10.1111/desc.12843


https://doi.org/10.1037/REV0000175


https://doi.org/10.1038/s41593-019-0500-6


https://doi.org/10.1523/JNEUROSCI.2770-10.2010


https://doi.org/10.1016/J.COBEHA.2020.05.014


Iordanou, K., Kuhn, D., Matos, F., Shi, Y., & Hemberger, L. (2019). Learning by arguing. Learning and Instruction, 63, 101207. https://doi.org/10.1016/j.learninstruc.2019.05.004


Kepecs, A., & Mainen, Z. F. (2012). A computational framework for the study


https://doi.org/10.3389/FNBOT.2013.00009

https://doi.org/10.1016/j.neuroimage.2004.10.024


https://doi.org/10.1186/S41235-017-0065-4


https://doi.org/10.1098/rstb.2012.0006

https://doi.org/10.31234/OSF.IO/TCD7P


https://doi.org/10.1111/PHPR.12145


Nuckolls, J., & Swanson, T. (2018). Respectable uncertainty and pathetic truth


1601 Pouget, A., Drugowitsch, J., & Kepecs, A. (2016). Confidence and certainty:
1602  distinct probabilistic quantities for different goals. Nature Neuroscience,
1604  Proust, Joelle. (2012). Metacognition and mindreading: one or two functions? In
1605  M. J. Beran, J. L. Brandl, J. Perner, & J. Proust (Eds.), Foundations of
1608  Metzinger & J. M. Windt (Eds.), Open MIND (Vol. 31, pp. 1–26). Johannes
1609  Gutenberg Universität Mainz. https://doi.org/10.15502/9783958570047
1610  Proust, Joëlle. (2019). From comparative studies to interdisciplinary research on
1612  https://doi.org/10.26451/abc.06.04.10.2019
1614  Press.
1616  explanation to improve mathematics learning: A meta-analysis and
1617  instructional design principles. ZDM - Mathematics Education, 49(4), 599–
1618  611. https://doi.org/10.1007/S11888-017-0834-Z
1621  doi.org/10.1016/J.TICS.2019.11.005
1623  anthropology and the opacity of other minds. Anthropological Quarterly,
1624  81(2), 407–420. https://doi.org/10.1353/ANQ.0.0005
1627  https://doi.org/10.3765/sp.5.6
1629  University Press.
1630  Rogoff, B., Coppens, A. D., Alcalá, L., Aceves-Azuara, I., Ruvalcaba, O.,
1633  https://doi.org/10.1177/1745691617718355
1634  Rogoff, B., Mistry, J., Göncü, A., Mosier, C., Chavajay, P., Heath, S. B., &
1635  Goncu, A. (1993). Guided Participation in Cultural Activity by Toddlers
1636  and Caregivers. Monographs of the Society for Research in Child
1637  Development, 58(8), i. https://doi.org/10.2307/1166109
1639  children develop an understanding of their own lack of knowledge. Child
1640  Development, 83, 1869–1883. https://doi.org/10.1111/j.1467-
1641  8624.2012.01830.x
1642  Rouault, M., & Koechlin, E. (2018). Prefrontal function and cognitive control:
1643  from action to language. Current Opinion in Behavioral Sciences, 21, 106–
1644  111. https://doi.org/10.1016/J.COBEHA.2018.03.008


