COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 47 (2023) e13253

ISSN: 1551-6709 online
DOI: 10.1111/cogs.13253

The Multifaceted Role of Self-Generated Question Asking in Curiosity-Driven Learning

Kara Kedrick,[a] Paul Schrater,[a,b] Wilma Koutstaal[a]

[a]Department of Psychology, University of Minnesota
[b]Department of Computer Science and Engineering, University of Minnesota

Received 19 April 2022; received in revised form 16 December 2022; accepted 30 January 2023

Abstract

Curiosity motivates the search for missing information, driving learning, scientific discovery, and innovation. Yet, identifying that there is a gap in one's knowledge is itself a critical step, and may demand that one formulate a question to precisely express what is missing. Our work captures the integral role of self-generated questions during the acquisition of new information, which we refer to as active-curiosity-driven learning. We tested active-curiosity-driven learning using our "Curiosity Question & Answer Task" paradigm, where participants (*N*=135) were asked to generate questions in response to novel, incomplete factual statements and provided the opportunity to forage for answers. We also introduce new measures of question quality that express how well questions capture stimulus and foraging information. We hypothesized that active question asking should influence behavior across the stages of our task by increasing the probability that participants express curiosity, forage for answers, and remember what they had thereby discovered. We found that individuals who asked a high number of quality questions experienced elevated curiosity, were more likely to pursue missing information that was semantically related to their questions, and more likely to retain the information on a later cued recall test. Additional analyses revealed that curiosity played a predominant role in motivating participants to forage for missing information, and that both curiosity and satisfaction with the acquired information boosted memory recall. Overall, our results suggest that asking questions

Correspondence should be sent to Kara Kedrick, Department of Psychology, University of Minnesota, N218 Elliott Hall, 75 East River Parkway, Minneapolis, MN 55455, USA. E-mail: kedri001@umn.edu

enhances the value of missing information, with important implications for learning and discovery of all forms.

Keywords: Question asking; Curiosity; Intrinsic motivation; Information foraging; Memory; Semantic analysis

1. Introduction

Curiosity is widely viewed as an intrinsically motivated state that drives learning and is often depicted within a broader learning context that begins with the detection of an information gap (Gruber & Ranganath, 2019; Kang et al., 2009; Loewenstein, 1994). But, the learning benefits of actively and autonomously generating questions about information gaps remains underexplored. Sometimes, the environment explicitly provides information gaps, such as in a classroom when a teacher poses a question (Dillon, 1988). Yet, many naturalistic settings present implicit information gaps that must be actively identified and translated into questions (Coenen et al., 2019). For example, we may recognize that an event has occurred (e.g., the dinosaurs became extinct) without a straightforward cause, leading us to query the world for answers (e.g., What caused the dinosaurs to become extinct?). The focus of much experimental work in cognitive psychology and neuroscience has been on the causes and course of curiosity about explicit information gaps, where participants are presented with gaps in the form of trivia questions (e.g., What instrument was invented to sound like a human singing?; Kang et al., 2009; Marvin & Shohamy, 2016). Failing to also consider situations where participants must pose their own questions creates an incomplete view of curiosity-driven learning. Asking quality questions about missing information is necessary for most real-world applications of curiosity because it sets the stage both for determining what information should be sought and for learning from information subsequently acquired (Golman & Loewenstein, 2018; Golman et al., 2021; Graesser & Person, 1994; Graesser & Olde, 2003).

In the work reported here, we extend the field's understanding of intrinsically-motivated learning by capturing the benefits of question-guided curiosity-driven learning. We substantiate our claims using an experimental (lab-based) paradigm in which participants are presented with a diverse range of existing facts that they could learn and integrate into their body of knowledge. Critically, the stimuli contained *implicit* information gaps—that is, gaps that indirectly implied to the participant that there was more to know about the given topic, pertaining to such information as what, how, or why. The gaps were not extrinsically provided in the stimulus or by an interlocutor, thereby prompting participants to autonomously generate their own questions. Notably, the paradigm was constructed such that we could assess participants' question-asking and curiosity from several further vantage points. After participants generated the questions that naturally arose for them about a given stimulus, they were asked to indicate how curious they were to learn more about the subject, and given the opportunity to forage for information that could bridge the gaps and extend their understanding of the topic. Finally, we assessed how well the participants encoded the acquired information through a cued-memory recall task. Our assessment also includes other factors that have been associated with curiosity-based learning. Specifically, we assessed how learning is impacted by the participants' interest or satisfaction with the information they found during foraging, their pre-experimental familiarity with the material, and whether the material related to their independently assessed broader individual interest.

1.1. Processes underlying active-curiosity-driven learning

Fig. 1 illustrates the processes that facilitate active-curiosity-driven learning, where we contextualize question generation among existing theories on curiosity-driven learning. We focus on the type of curiosity, where an individual is curious to know information that pertains to a specific information gap (Berlyne, 1960; Litman, 2008; Loewenstein, 1994). The identification of a specific information gap allows an individual to translate the missing information into a question and then query the environment for answers (Golman & Loewenstein, 2018; Golman et al., 2021). Gap-focused curiosity contrasts with more wide-ranging diversive forms of curiosity where one is curious to know more information without having a particular gap in mind (Berlyne, 1960; Litman, 2008). Fig. 1 demonstrates the connection between our lab-based task and active-curiosity-driven-learning, which encompasses the internal process triggered by each phase of the task and the outcome of the internal process.

Unlike contexts where information gaps are explicitly provided, active-curiosity-driven-learning starts with the initial representation of new information. The generation of a mental representation is facilitated by parsing the incoming information into entities and relationships and using prior knowledge to further develop the representation ("Represent New Information" in Internal Process, Fig. 1; Johnson-Laird, 1983; Kintsch, 1988). By actively representing the new information, a person develops both a positive understanding of what is there, as well as an account of the gaps in their representation. Developing a clear, well-formed mental representation allows for the detection of information gaps ("Detect Gap" in Internal Process; Fig. 1). Detected gaps can be translated into questions that guide subsequent search for answers ("Question Asking" in Outcome of Internal Process; Fig. 1). The quality of questions asked depends on how well the new information has been represented, wherein accurate representations invite questions of higher quality that better capture information gaps. After a gap has been detected, gaps are appraised or evaluated based on their uncertainty or familiarity (Gruber & Ranganath, 2019; Kang et al., 2009), prior interest or appreciation for related material (Hidi & Renninger, 2006; Sharot & Sunstein, 2020), and other measures of value ("Appraise Gap" in Internal Process and "Familiarity & Individual Interest" in Outcome of Internal Process; Fig. 1). This evaluation process likely involves multiple intrinsic and extrinsic factors (Sharot & Sunstein, 2020), which contribute to the urgency with which individuals would actively seek information for an identified gap.

The outcome of the gap appraisal process determines whether or not a person is curious to fill the identified gap made salient by the evaluation process ("Curiosity" in Outcome of Internal Process; Fig. 1), and reflects the expected value associated with acquiring the missing information (Gruber & Ranganath, 2019; Kang et al., 2009; Loewenstein, 1994). If gap appraisal results in curiosity, then the person is driven to forage for or seek out the missing

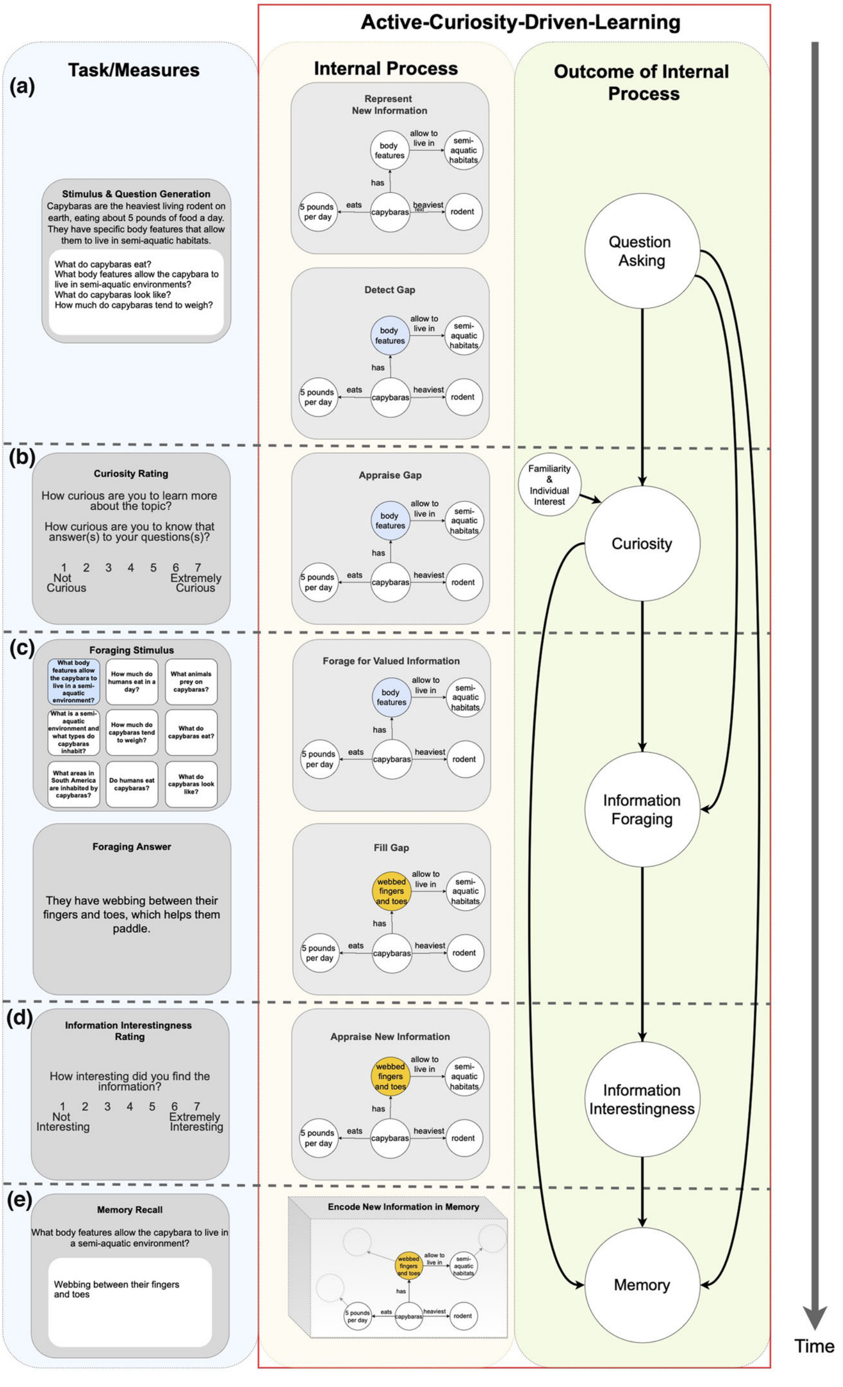

Fig. 1. A schematic summary of active-curiosity-driven learning, represented in the "Internal Process" column and the "Outcome of Internal Process" column. The sequence of steps in the Curiosity Q&A Task that correspond to each part of the learning process are represented in the "Task/Measures" column. (a) This panel depicts what

happens when someone first encounters new information. Specifically, an individual can actively create a mental representation by identifying the salient entities and relationships between those entities in the information that is given. Next, the representation can be used to detect gaps within the information that was provided, and those gaps can be translated into questions. In the Q&A task, participants are presented with factual statements that must be actively represented in order to identify implicit information gaps. We measured whether the implicit information gaps were detected by having participants type questions they had about the stimulus (example possible participant-generated questions are shown in the white text box). (b) After an information gap has been identified, it is appraised or evaluated, with the potential for the appraisal process to result in curiosity. In the task, we asked participants to provide curiosity ratings as a measure of their desire to know more about their questions and the topic. (c) If the gap sparks curiosity, the individual may choose to forage for valued information or search for other gaps that can be translated into questions, the former of which is tested in our study. In the task, participants were presented with nine "buttons" that displayed different questions about the gaps in the factual statement as well as topic-related questions. Participants could choose to forage for information by pressing the button to see the answer to any of the questions. Once the information has been acquired, the gap within the individual's mental representation can be filled in with the missing information. Upon selecting the gap question, participants were provided with the gap answer that could be used to fill the gap within their mental representation. (d) After the new information has been acquired, it is appraised or evaluated. In the task, we measured the value of the new information by asking participants to report how interesting they found the new information. (e) After the new information is appraised, the mental representation is encoded in memory. In the task, we tested participants' memory by asking them to recall the answers to the gap questions (example of a participant's memory-based recall shown in the white text box).

information ("Information Foraging" in Outcome of Internal Process, Fig. 1; Gruber & Ranganath, 2019). When the missing information has been acquired, it can be used to fill the information gap ("Fill Gap" in Internal Process; Fig. 1). The acquired information is then further evaluated to assess how interesting or satisfying the missing information is and/or how well it fills the information gap ("Appraise New Information" in Internal Process and "Information Interestingness" in Outcome of Internal Process, Fig. 1; Marvin & Shohamy, 2016; McGillivray, Murayama, & Castel, 2015). Finally, the representation is encoded ("Encode New Information in Memory" in Internal Process and "Memory" in Outcome of Internal Process; Fig. 1). The encoding process has been shown to depend on both curiosity (the predicted value of the missing information; Gruber, Gelman, & Ranganath, 2014; Kang et al., 2009) and interestingness or satisfaction (the actual value of the information; Marvin & Shohamy, 2016; McGillivray, Murayama, & Castel, 2015).

Beyond these situationally related or state-based responses to encountering novel or incomplete information, active curiosity may also be influenced by individual differences, such as someone's general predispositional "trait" curiosity and their openness to experience (Litman, 2008; Woo et al., 2013). Trait curiosity has been differentiated into a form in which someone is aware of what they do not know (or gaps in their knowledge) and is motivated by a desire to escape ignorance or uncertainty, called Deprivation-Type Curiosity, and a more diffuse form in which one generally enjoys learning something new, without focusing on specific information gaps, called Interest-Type Curiosity (Litman, 2008). In the context of active-curiosity-driven learning, the focus of Deprivation-Type Curiosity on resolving information gaps makes it more relevant than Interest-Type Curiosity. Openness to experience is often characterized as an individuals' proclivity toward novel and variable phenomenon and experiences, including the motivated cognitive exploration of novel and varied information (DeYoung et al., 2005; Woo et al., 2013). Given that both individual differences in curiosity and openness to experience may influence active-curiosity-driven learning, they were also included in our analyses.

Our goal is to understand the role that active question asking and curiosity play in learning, by: (1) unpacking curiosity-driven-learning into components and integrating active question asking; (2) introducing new measures sensitive to these components; and (3) contextualizing and relating the different measures of learning motivation to provide a clearer picture of the role of different components within the overall learning process. Active-curiosity-driven learning involves both the internal processing of information and the outcome of those internal processes. The outcome is comparatively easier to test than the internal processes; thus, our lab-based task focuses on capturing the outcome, and provides indirect evidence for the internal processes. As suggested by prior literature, we suspected that active-curiosity-driven learning would be influenced by both state differences (changes in response influenced by stimuli) and individual differences (differences in aggregated responses depending on the participant).

1.2. Question asking, information seeking, and memory

The pivotal role of actively generating questions about implicit information gaps is supported by evidence pointing to the benefits of self-directed learning—particularly providing learners with the opportunity to make their own decisions or choices about the information they want to experience. For example, self-directed learning may allow the individual to focus their search efforts on information they do not yet know, prompt them to uncover information that would not be noticed under more passive learning conditions (Gureckis & Markant, 2012), or enable them to adjust the pacing of acquiring new information to match their own attentional or motivational state (Markant et al., 2014, 2016; Markant, 2020). The higher degree of mental effort associated with self-directed learning may elicit more careful consideration of what specific information should be sought out, constituting a form of "desirable difficulty" that leads to more appropriate inquiry and the asking of more informative questions (Kachergis et al., 2017). Asking and pursuing questions is an instance of self-directed learning, where a person's search for answers is guided by the gaps they inquire about in their mental representation. But not all questions are equivalent. Rather, the quality of a question is the product of a high-quality mental representation and should result in the acquisition of information of higher relevancy.

Question asking not only has the potential to provide a more directed information-seeking trajectory, but may also facilitate encoding of the newly acquired (previously missing) information. Support for this is based on evidence for the beneficial memory consequences arising from self-generating material or "the generation effect" (Bertsch et al., 2007; Rosner et al., 2013; Slamecka & Graf, 1978) and active, self-directed learning (Gureckis & Markant, 2012; Kachergis et al., 2017). The clarity of the initial representation–that facilitated the quality and quantity of questions generated–may also make it easier to update that representation once the information has been acquired, enhancing the encoding process. Question generation will also activate relevant surrounding information, contributing to more effective storage of the newly encountered information (Jirout, 2020; Tawfik et al., 2020).

1.3. Hypotheses and description of study

Guided by the above-mentioned empirical and theoretical work, we tested the predictors of three outcomes that are typically associated with curiosity-driven learning, articulated in the following hypotheses:

1. Curiosity is positively associated with the quality and quantity of question asking and the expected value of receiving the gap-related information, measured by familiarity and individual interest.
2. Information foraging is enhanced by asking quality questions about foraged-for information and curiosity to know more.
3. Memory recall will be boosted by the quality and quantity of question asking, curiosity to know more, and interest in or satisfaction with the new information.

We test our hypotheses in terms of (1) state-based fluctuations within participants' responses and (2) individual differences between participants. For the individual difference analyses, we also included trait measures of curiosity and openness to experience as predictors of curiosity, information foraging, and memory.

We manipulate the participants' investment in active gap detection by introducing a question asking paradigm called the Curiosity Question & Answer (Q&A) Task. This task allows for the examination of what causes individuals to forage for and remember external information gaps (information that was implicitly missing from the provided statement) and subjective-information gaps (information that participants' deem as missing given their prior knowledge). To emulate real-world instances of information foraging, our stimuli were based on a diverse range of existing facts that participants could learn and integrate into their body of knowledge. The stimuli included multiple information gaps, where the text hinted that there was more to know about a given topic. After reading the information, participants were asked to generate related questions, which may or may not address the information gaps within the stimulus. The questions asked were assessed in terms of quantity and quality. Quantity was assessed by counting the number of questions asked. Quality was assessed using a computational measure of text similarity (described later) that compared, first, the similarity of the text of participants' questions to that of the presented stimuli, and second, the similarity of the text of participants' questions to that of the questions presented in the information foraging phase, that participants could choose to look at and read during that phase. Afterward, participants rated their curiosity to know the answers to their questions and to know more about the topic. Next, participants had the opportunity to forage for missing information by choosing to see answers to questions that relate to the gaps in the stimulus and to the general topic. This broadly mimicked, in the lab, how individuals might search for and encounter new information on the Web or in a document, with newly found information itself prompting further search (Pirolli, 2005; Pirolli & Card, 1999; Savolainen, 2018). After foraging, participants were asked to rate how interesting they found the new information. Finally, we administered a surprise cued-recall memory task to assess what factors shape later memory for gap-related information.

2. Method

2.1. Participants

A total of 140 participants took part in the experiment. The results for five participants were omitted: one due to technical issues, two as a result of incorrect administration of the paradigm, one because they did not produce any behaviors in the information foraging phase, and one because they did not fill out the Deprivation-Type Curiosity scale.

The final set of participants included 135 undergraduates from a large midwestern university (97 females, *M age* = 19.93 years, *SD* = 2.39). All participants were required to be within the age range of 18–30 years old, to have learned English by the age of 6 years, and to have normal (or corrected) hearing and vision. In exchange for their participation, participants were awarded extra course credit points. Our research protocol was approved by the Institutional Review Board.

2.2. Materials

2.2.1. Curiosity Q&A Task

In the Curiosity Q&A Task, participants were presented with 18 statements about a variety of obscure, nontechnical topics (provided in Appendix S1). All statements contained one to three sentences, and on average included 28.11 words (*SD* = 3.41; range = 21–33 words). The 18 statements were listed in two distinct orders (Set A and Set B). Here is an example statement:

> *Capybaras are the heaviest living rodent on earth, eating about 5 pounds of food a day. They have specific body features that allow them to live in semi-aquatic habitats in South America.*

Each of the statements included two information gaps. These gaps were meant to imply to the participant that there was more to know about the given topic, pertaining to such information as who, what, when, how, where, and why. Therefore, a clear question was associated with each gap (e.g., for the above statement, one gap-related question is, “What body features allow the Capybara to live in a semi-aquatic environment?”).

The “Task/Measures” section of Fig. 1 presents a schematic overview of the task’s format. For the Curiosity Q&A Task, participants read each statement on a computer monitor and were instructed to type any questions they have after reading the statement. They were informed that they would have the opportunity to view the answers to some of their questions, and asked to write “Pass” if they do not have any questions. Participants were given at least 20 seconds to both read the stimulus and write questions before they could progress to the next page. There was no limit on the number of questions they could type. After typing all their questions, participants were prompted to provide three ratings, each given on a 7-point Likert scale (1 = not at all; 7 = extremely): how curious they were to learn more about the statement-related topic, how curious they were to know the answer(s) to their question(s), and how familiar they were with the factual information beforehand. If the participant did not type any questions for a given statement, they were instructed to leave the second rating blank. To avoid revealing the gaps in the stimulus, we asked participants to rate their curiosity to know all of the answers to their questions. Our measures of curiosity to learn more about the stimulus-related topic and to know the answers to their questions allowed us to see how curiosity, with or without the explicit identification of the missing information, promotes gap-directed information seeking.

Next, participants were presented with nine "buttons" that displayed topic-related questions, allowing for information-foraging opportunities. Two of the buttons presented gap-related questions. The other seven buttons displayed questions that related to the topic, and (as described later in this section) were comprised of commonly asked questions from our pilot study. Participants had the opportunity to click on any of the buttons to look at the answer to the question. There was no time limit nor a restriction on the number of buttons they could press. We used a counterbalancing scheme such that, across all the factual stimuli, the gap-related questions were about equally likely to appear on any one of the nine "buttons," and had two different question orders for Set A and Set B. When participants finished information foraging for a given factual statement, they were asked to rate how interested they were in the answers they had looked at on a 7-point Likert scale. Once they completed their information interestingness rating, the entire process repeated for the next factual statement, continuing until all 18 factual statements were presented.

In pilot work, with 26 participants drawn from a similar population as tested here and using an auditory-based presentation, participants on average generated 1.53 ($SD = 0.72$) of the gap-related questions for each stimulus. In this task, an experimenter read each stimulus to the participant. Afterward, the participant asked questions, and the experimenter responded with scripted answers. This provided assurance that the gaps were detectable by the participants. It also allowed us to collect commonly asked questions— that did not pertain to the stimulus gaps—for use as alternative questions in the foraging phase of our main experiment.

2.2.2. *Question asking measures*

We created measures that assessed both the quantity and quality of questions asked. Quantity denotes the number of questions posed, but does not indicate how relevant the questions are to the stimulus. Quality indicates how well the information mentioned in the questions maps onto the stimulus, and indirectly speaks to the quality of the participants' mental representation (i.e., ability to accurately capture the stimulus text and what was implicitly missing from the stimulus).

In terms of quantity, we measured (1) the number of questions asked that directly pertained to the information gaps within the stimulus (referred to as Stimulus-Gap Questions) and (2) the number of questions asked about topic-related information missing from the participants' prior knowledge (referred to as Subjective-Gap Questions). For example, let us consider the scores received if a participant provided the following questions in response to the Capybara stimulus (i.e., Capybaras are the heaviest living rodent on earth, eating about 5 pounds of food a day. They have specific body features that allow them to live in semi-aquatic habitats in South America):

Question 1. *"What body features allow Capybaras to live in a semi-aquatic environment?"*

Question 2. *"What do Capybaras eat?"*

The participant would have received one point toward their Stimulus-Gap Questions' score because Question 1 directly targets one of the implicit gaps within the stimulus. The participant would also receive one point toward their Subjective-Gap Questions' score because Question 2 does not pertain to the implicit information gaps within the stimulus, but does pertain to that individual's own understanding of Capybaras more generally—including that individual's recognition of additional information they might request, so as to expand or better connect their knowledge regarding those creatures.

For the Stimulus-Gap Questions, two independent raters scored the questions and showed high reliability ($r = 0.83$). A third rater identified and resolved any divergences between the two raters. For a small subset of responses, the typed questions provided by the participant only partially or ambiguously identified the gap question, and so were awarded a score between 0 and 1 (i.e., either 0.5 or 0.75). Given that these scores were infrequent (91 instances out of 2,430), they were removed from the analyses. The maximum Stimulus-Gap Questions score per stimulus was two, indicating that the participant asked both gap questions for a given factual stimulus.

We measured the quality of questions asked from multiple perspectives. First, we assessed the textual similarity between participants' questions and the stimulus text, where high similarity indicates that participants' questions did a better job of capturing information conveyed in the stimulus than questions that are low in similarity. Second, we assessed the similarity between the text from the group of questions the participants typed for each stimulus and the text from the group of questions that participants could see the answers to during the corresponding foraging phase. Third, we assessed the similarity between the text from a specific question a participant typed and the text from a specific question in the foraging phase. Measuring the semantic similarity between the participants' questions and the foraging questions allows us to see whether participants' foraging behavior was influenced by how well their questions related to the information they could uncover during the foraging phase.

To measure semantic similarity, we used term-frequency inverse-document frequency (tf-idf) and cosine similarity. Tf-idf is a common machine learning technique that indicates which words distinguish documents from a set of documents, by identifying highly informative or semantically relevant words and down-weighting common words (Salton & Buckley, 1988). The process generates a vector of tf-idf scores for each document that correspond to all unique words from a set of documents. Cosine similarity computes the semantic similarity between documents by assessing the congruence of pairs of tf-idf vectors. Quantifying the semantic similarity between pairs of documents using the cosine similarity of tf-idf vectors has been used to study a variety of topics, such as the similarity between text from different Wikipedia pages visited during information seeking (Lydon-Staley et al., 2021), to predict citations between pairs of research articles (Shibata et al., 2012), and to recommend research articles based on user queries (Philip et al., 2014). See Appendix S4 for more information.

First, to better understand how each participants' questions related to the stimulus, we compared the text found in the participants' questions to the text found in the stimulus. We began by calculating tf-idf for the participants' questions and stimulus text. For the participants' questions, our unit of analyses was the group of questions each participant typed in response to a given stimulus, leaving us with 18 sets of questions per participant. Our entire corpus included the participants' grouped questions and the stimulus text. Then, we found the semantic similarity between the participants' questions and the corresponding stimulus by computing the cosine similarity between the questions' tf-idf vectors and corresponding stimulus tf-idf vectors (referred to as Semantic Similarity: Questions-to-Stimuli).

Second, to quantify the similarity between the participants' questions and the questions they could seek answers to during the information foraging phase, we compared the text from the participants' questions to the text from the questions in the information foraging phase. We calculated tf-idf for the participants' questions and foraging questions. Our unit of analyses for the participants' questions is the same as above. For the foraging questions, each document included all of the per stimulus text from the questions associated with the nine buttons in the information foraging phase. Our entire corpus included the participants' grouped questions and the text from the grouped-foraging questions. Afterward, we found the semantic similarity between the participants' questions and the corresponding foraging questions by computing the cosine similarity between the participants questions' tf-idf vectors and foraging questions' tf-idf vectors (referred to as Semantic Similarity: Grouped Questions-to-Foraging Questions).

Third, to more precisely quantify the similarity between a specific question asked by the participant and a specific question from the foraging phase, we compared the text from each question typed by the participant to the text from each question from the foraging phase. We calculated tf-idf for the participants' questions and foraging questions. For the participants' questions, our unit of analyses was each individual question the participant typed in response to a given stimulus. For the foraging questions, each document included the text from one of the questions listed on the nine buttons in the information foraging phase. Our entire corpus included the participants' individual questions and the text from each foraging question. Afterward, we found the semantic similarity between each of the participants' questions and each of the foraging questions by computing the cosine similarity between the participant questions' tf-idf vectors and corresponding foraging questions' tf-idf vectors (referred to as Semantic Similarity: Single Question-to-Foraging Question). To focus our analysis, we subset the data to only include the highest cosine similarity for each foraging question. Thus, we were left with the semantic similarity associated with the participant's question that was the best match for the foraging question.

At a broad conceptual level, our question-asking measures are indicative of how well the participant has represented the new information, allowing for active gap detection. That is, both the quantity (number of questions) and the quality (how well the questions capture the context) depend on one's mental representation. Therefore, the measures are not meant to be independent. As expected, the question-asking measures were highly correlated, especially when measures were aggregated for the individual difference analyses (r between 0.39 and 0.90). Including measures that are highly correlated poses multicollinearity issues while running regression models. Therefore, we sought to decouple the question-asking predictors, making the measures independent, by running multiple principal component analyses (PCA; prcomp, stats package; R Core Team, 2013). First, we ran PCA on the subset of question-asking predictors used to predict curiosity, namely, the Stimulus-Gap Questions, Subjective-Gap Questions, and Semantic Similarity: Questions-to-Stimuli. We did not include the Semantic Similarity: Grouped Questions-to-Foraging Questions or the Semantic Similarity: Single Question-to-Foraging Question measures because the foraging phase (where participants could choose to look at the answers) occurred after the participant reported their state curiosity, making it irrelevant to their state curiosity judgment. All three principal components were used as predictors in our regression models.

In addition, we ran PCA using the question-asking predictors for our information foraging and memory regression analyses, which included Stimulus-Gap Questions, Subjective-Gap Questions, Semantic Similarity: Questions-to-Stimuli, and Semantic Similarity: Grouped Questions-to-Foraging Questions. Unlike for our curiosity analyses, we included Semantic Similarity: Grouped Questions-to-Foraging Questions because the participants' foraging happens concurrently with the foraging behavior and prior to the memory task. We did not include Semantic Similarity: Single Question-to-Foraging Question because our foraging and memory models involve per-stimulus responses, whereas this measure accounts for the responses to each of the nine questions in the foraging phase. The Semantic Similarity: Single Question-to-Foraging Question measure is used for a separate analysis detailed in the "Impact of state-based changes on learning" section. All four principal components were used as predictors in our regression models.

2.2.3. Curiosity and information interestingness measures

Our curiosity measure was the sum of the self-reported ratings for our two curiosity questions, each reported using a 7-point Likert scale: 1) curiosity about the participant's generated questions and 2) curiosity to know more about the topic. In cases where participants did not type any questions, they were assigned a score of 0 and asked to report their curiosity to know more about the topic. Adding both of these measures together allowed us to capture participants' general curiosity to know more, even if the participant did not detect gaps by asking questions. The combined curiosity measure was used as the target variable for the curiosity regression models, and was predicted by question-asking behaviors. In order to use curiosity as a stand alone predictor of foraging and memory, we removed the influence of question asking from our curiosity measure (referred to as residual curiosity). To do this, we ran a linear mixed model for the state analyses, wherein we only included the question-asking principal components that significantly predicted curiosity (i.e., principal component 2 and principal component 3 described in more detail in the "Impact of state-based changes on learning" section). We also ran a linear model for the individual-difference analyses, wherein we only included the question-asking principal components that significantly predicted curiosity (i.e., principal component 1 and principal component 3, described in more detail in the "Impact of individual differences on learning" section). Then, we subtracted the amount of curiosity predicted by question asking (predict function; Fox & Weisberg, 2019) from our curiosity measures, separately for the state analyses and the individual difference analyses.

We also modified our information interestingness measure so that it could be a stand alone predictor of memory, by removing the influence of curiosity (referred to as residual information interestingness). To do this, we ran a linear mixed model for the state analyses and a linear model for the individual-difference analyses, wherein we only included residual curiosity as a predictor. Then, we subtracted the amount of information interestingness predicted by residual curiosity (predict function; Fox & Weisberg, 2019) from our information interestingness measures, separately for the state analyses and the individual difference analyses.

2.2.4. Information foraging measures

We calculated multiple measures of information foraging: (1) the number of times a participant looked at the stimulus-gap answers (referred to as Stimulus-Gap Foraging); (2) the number of times the participant looked at the answers to questions that were not related to the stimulus gaps (referred to as Subjective-Gap Foraging); and (3) whether the participant looked at the answer to a specific question that could either be related to or not related to the stimulus gap (referred to as Single-Question Foraging). For Stimulus-Gap Foraging, participants were awarded one point per gap answer viewed, for a maximum of two points per stimulus. For Subjective-Gap Foraging, participants were awarded one point per answer viewed, for a maximum of seven points per stimulus. For Single-Question Foraging, participants were awarded one point if they chose to look at an answer and zero points otherwise. Multiple viewings of the same answer were not counted toward their scores.

2.2.5. Pre-experimental individual interest inventory

In the current study, we examined whether a person's individual interest— topics that were preidentified by one as being of interest—would influence curiosity. The Interest Inventory, newly developed as part of this project, was administered separately from the Q&A task and without any reference to that task (see Appendix S2). It included 94 items, in 12 broad domains (e.g., Performing Arts, Social Sciences, and Games/Sports), and participants rated their interest in each item on a 7-point Likert scale. The primary and secondary relevance of each of these interests to the factual stimuli of the Q&A task was determined by consensus discussion with six raters. The analyses reported here averaged the primary and secondary relevance ratings for each stimulus into a composite assessment of participants' pre-experimental or individual interest regarding each factual stimulus.

2.2.6. Memory task

In the cued-recall memory phase, participants provided typed answers to questions about the information they had earlier encountered during the Curiosity Q&A Task. They were asked to answer questions that pertained to the two information gaps within each factual statement, for a total of 36 cued-recall questions. The maximum score was 2 per stimulus, which was scaled between 0 and 100 for analysis (discussed in Appendix S6). For example, as it relates to the Capybara stimulus shown in Fig. 1, participants would see the question, "What body features allow the capybara to live in a semi-aquatic environment?" and be asked to type the answer. Each complete and correct answer was awarded a total of one point. In order to score the memory task, each answer was divided into substantive components, that is, aspects of the answer that entailed important or detailed information. For example, the answer to the above-mentioned question is: "They have webbing between their fingers and toes, which helps them paddle, and eyes and ears positioned high on their heads, which allows them to hide well underwater." Participants would receive one point if their answers included any of the following: webbing, fingers, toes, eyes, ears, and high on their heads. The memory score per stimulus was the proportion of items the participant got correct. Two independent raters scored the participants' answers and had high inter-rater reliability ($r = 0.98$).

2.2.7. *Trait measures*

We were also interested in testing whether traits—that are commonly associated with curiosity—predicted behaviors. Therefore, we include measures of trait curiosity and openness to experience in our analyses. To measure trait curiosity, we administered the Deprivation-Type Epistemic Curiosity Scale (Litman, 2008), which includes five items (Cronbach's alpha = 0.82). Participants were asked to indicate how much they agreed with each item on a 4-point Likert scale, where 1 indicates "Almost Never" and 4 indicates "Almost Always."

To measure openness to experience, we administered the well-validated Big Five Aspect Scales (BFAS; DeYoung, Quilty, & Peterson, 2007): a 100-item questionnaire that assesses the big five personality dimensions, and separates each dimension into two aspects (10 items per aspect). We focused on the aspects of openness to experience (i.e., Openness and Intellect). The Openness aspect concerns one's appreciation for imaginative activities or aesthetics (Cronbach's alpha = 0.77), whereas the Intellect aspect captures one's perceived ingenuity and proclivity for analytical thinking (Cronbach's alpha = 0.81). For this task, participants rated the degree to which they found each statement to be self-descriptive on a 5-point Likert scale, where 1 indicates "Strongly Disagree" and 5 indicates "Strongly Agree."

We also included the Openness to Experience scale (Woo et al., 2013) that differentiates the two commonly held intermediate factors of openness—intellect and culture—into six subscales, nine items each. We only included the intellect factor subscale of "Curiosity" in our analyses, which taps one's desire to learn about scientific and intellectual information (Cronbach's alpha = 0.69). Participants rated how much they agreed with each statement on a 5-point Likert scale, where 1 indicates "Strongly Disagree" and 5 indicates "Strongly Agree."

2.3. *Procedure*

After providing informed consent, the participants took the Curiosity Q&A Task. The task was administered on a desktop computer via E-Prime 3 (Psychology Software Tools). It began with a practice stimulus to give participants a sense of how the task would proceed, and an opportunity to ask for any clarifications. Depending on the individual participant's responses, the Q&A Task lasted between 30 min and 1 h. Participants were block randomized to receive the stimuli in the order of either Set A or Set B.

After they finished the Q&A Task, participants completed the Pre-Experimental Interest Inventory as well as self-report state and trait questionnaires of openness to experience and curiosity (see Appendix S3 for additional measure of Tolerance of Ambiguity). Next, they were administered the cued-recall memory task. The order of the memory questions matched the order in which the stimuli were seen in either Set A or Set B, which allowed us to control for the retention interval. The questionnaires and the memory task took approximately 1 h to complete; participants were required to take one or more breaks. Finally, the participants were thanked, debriefed, and compensated.

2.4. Overview of analyses

Fig. 2 provides an overview of our tested models, and specifies our hypothesized predictors of curiosity, information foraging, and memory. Specifically, we analyzed: (1) whether *self-reported state curiosity* is predicted by question asking, familiarity, and individual interest; (2) whether *foraging for missing information* is predicted by question asking and curiosity; and (3) whether *memory recall of the acquired information* is predicted by question asking, curiosity, and information interestingness. We used regression analyses to test each hypothesis both in terms of state-based fluctuations (referred to as state analyses) and individual differences (referred to as individual difference analyses). For our state analyses, we scaled the repeated measures (for each of the 18 stimuli from the Curiosity Q&A Task) within-subjects; for our individual difference analyses, we averaged the repeated measures from the Curiosity Q&A Task for each participant and then scaled the measures between subjects. Testing both perspectives allowed us to see whether active-curiosity-driven-learning is modulated by both state and individual differences. In addition, the individual difference analyses also permitted the incorporation of personality predictors that are not influenced by the repeated measures in our task. That is, we were able to test whether personality trait measures of trait curiosity and openness to experience predict curiosity, information foraging, and memory.

3. Results

3.1. Impact of state-based changes on learning

We began by analyzing all responses at the state level, wherein variables were standardized within subjects. To test our hypotheses, we ran regression models for state curiosity, information foraging, and memory. Although curiosity ratings differed significantly depending on the stimulus order ($t = -2.68$, $p = .008$), the curiosity regression model was unable to converge with stimulus order as an additional random effect. Therefore, we excluded the random effect of stimulus order from the analyses. The stimulus order for Set A and Set B were not significantly different for information foraging ($t = 0.34$, $p = .73$) and memory ($t = -0.82$, $p = .41$), and thus were not included as random effects. We identified the best fitting models with the lowest BIC score using best subset selection, wherein we ran models for all combinations of predictors. If multiple models were not statistically significant (i.e., less than a 2-point difference in BIC score; Raftery, 1995), then we selected the model with

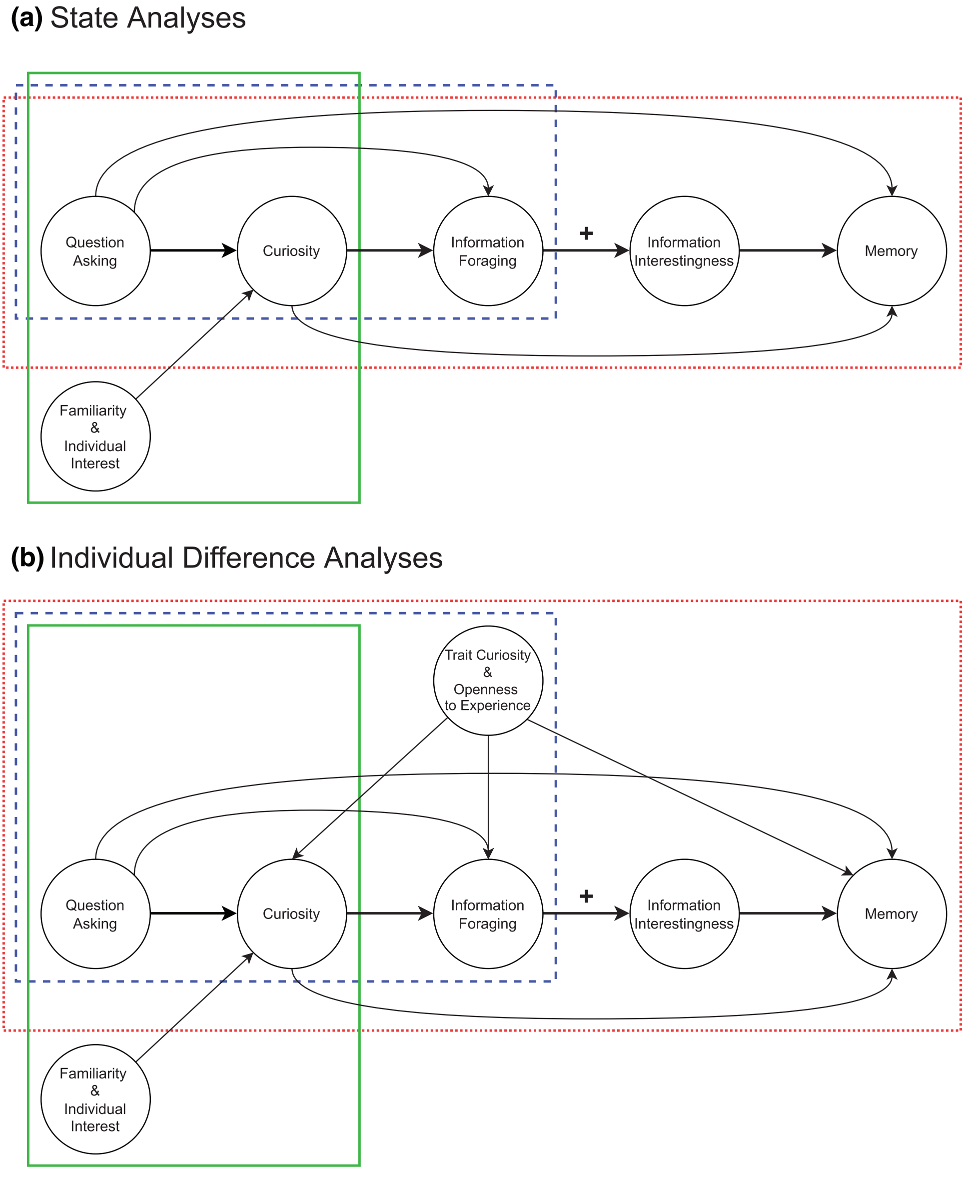

Fig. 2. Schematic of curiosity, information foraging, and memory regression models. (a) Depicts our model for the state analyses: (1) the green-solid line includes our hypothesis that self-reported state curiosity is predicted by question asking, familiarity, and individual interest; (2) the blue-dashed line includes our hypotheses that foraging for the missing information is predicted by question asking and state curiosity; and (3) the red-dotted line includes our hypotheses that memory of the acquired information is predicted by question asking, state curiosity, and information interestingness. The "+" above the arrow exiting "Information Foraging" indicates that our model of memory only included instances where participants viewed one or more of the gap answers during the information foraging phase. (b) Depicts our model for the individual difference analyses for curiosity, information foraging, and memory. The analyses include the same hypothesized predictors as those listed for the state analyses with the addition of trait curiosity and openness to experience.

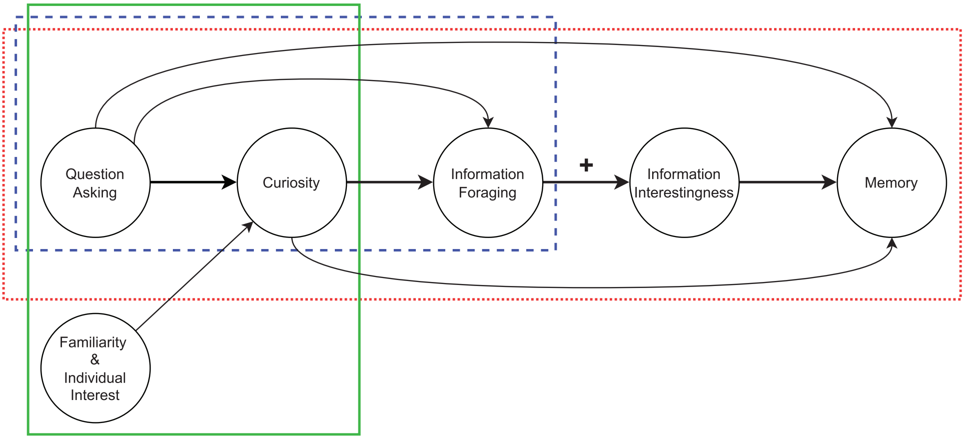

Fig. 3. State models of curiosity, information foraging, and memory. All hypothesized predictors were present in the best fitting models.

the fewest parameters. For curiosity, Subjective-Gap Foraging and memory, we ran Linear Mixed-Effects Models (lmer, lme4 package; Bates, Mächler, Bolker, & Walker, 2015) using RStudio (RStudio Team, 2020) with participant and stimulus as random additive effects. After selecting the best model for each analyses, we checked that each model met the assumptions for linear mixed effects models (see Appendix S6 for details). We also ran additional memory analyses (1) where we controlled for familiarity (see Appendix S7) and (2) where we ran a negative binomial Generalized Linear Mixed-Effects Model (see Appendix S6). For Stimulus-Gap Foraging, we ran a Poisson Generalized Linear Mixed-Effects Model (glmer, lme4 package; Bates, Mächler, Bolker, & Walker, 2015) using RStudio (RStudio Team, 2020) with participant and stimulus as random additive effects. We ran a Poisson Generalized Linear Mixed-Effects Model because our outcome variables are counts of positive integers and follow a Poisson distribution, wherein the mean and variance are approximately the same ($mean = 1.37$, $sd = 0.66$). For Single-Question Foraging, we ran a Logistic Generalized Linear Mixed-Effects Model (glmer, lme4 package; Bates, Mächler, Bolker, & Walker, 2015) using RStudio (RStudio Team, 2020) with participant and stimulus as random additive effects. We ran a Logistic Regression because our dependent variable was binary (i.e., 1 if the participant chose to look at the answer and 0 if they did not). See Table 1 for means, standard deviations, and correlations; and see Fig. 3 for a schematic overview of the results.

First, we analyzed our model of curiosity. We hypothesized that curiosity is associated with the efficiency of intrinsic gap detection (as measured by question asking) and the valuation of potentially receiving information pertinent to the detected gaps (i.e., expected value). This leads to the prediction that state curiosity judgments should be positively correlated with both question asking and measures of the expected value of addressing an information gap (as measured by familiarity and individual interest with a topic). As hypothesized, state curiosity was significantly predicted by question asking, familiarity, and individual interest, with a BIC score of 9,052.41 (see Table 2). Specifically, state curiosity is significantly

Table 1
Means, standard deviations, and correlations for the state analyses

	M	SD	1	2	3	4	5	6	7	8	9	10	11
1. Stimulus-Gap Questions	1.07	0.69	1.00										
2. Subjective-Gap Questions	1.53	1.85	–0.39	1.00									
3. Semantic Similarity: Q-to-Stim	0.29	0.15	0.37	–0.09	1.00								
4. Semantic Similarity: GQ-to-FQ	0.29	0.14	0.21	0.04	0.55	1.00							
5. Curiosity	8.04	2.96	0.13	0.23	0.05	0.02	1.00						
6. Familiarity	1.64	1.15	–0.02	0.00	–0.02	–0.03	0.26	1.00					
7. Individual Interest	4.12	1.31	–0.03	0.11	–0.03	0.00	0.23	0.14	1.00				
8. Stimulus-Gap Foraging	1.37	0.66	0.23	–0.06	0.14	0.06	0.11	0.01	0.03	1.00			
9. Subjective-Gap Foraging	2.92	2.21	–0.06	0.18	0.00	0.04	0.21	0.11	0.08	0.36	1.00		
10. Information Interestingness	4.3	1.66	0.12	0.15	0.00	0.01	0.60	0.22	0.24	0.21	0.30	1.00	
11. Memory	24.81	21.90	0.01	0.06	–0.04	0.05	0.13	–0.01	0.07	0.44	0.30	0.23	1.00

Note. Variables used for the state models of curiosity and information foraging. M and SD are used to represent mean and standard deviation, respectively (total of 1,986 instances and 132 participants). The ranges for the measures are as follows: Stimulus-Gap Questions (0-2), Subjective-Gap Questions (there was no cap on the number of questions), Semantic Similarity: Questions-to-Stimuli (Q-to-Stim; 0-1), Semantic Similarity: Grouped Questions-to-Foraging Questions (GQ-to-FQ; 0-1), Curiosity (1-14)-combined measure of curiosity about questions (0-7) and curiosity about topic (1-7), Familiarity (1-7), Average Individual Interest (1-7), Stimulus-Gap Foraging (0-2), Subjective-Gap Foraging (0-7), Information Interestingness (1-7), and Memory (0-100).

Table 2
Regression coefficients predicting state curiosity from question asking, familiarity, and individual interest

	Curiosity
(Intercept)	7.92***
	(0.20)
Question-Asking Principal Component 2	0.55***
	(0.06)
Question-Asking Principal Component 3	0.69***
	(0.08)
Familiarity	0.60***
	(0.06)
Individual Interest	0.42***
	(0.06)
Participant random effects	Y
Stimulus random effects	Y
AIC	9007.78
BIC	9052.41
Num. obs.	1958
Participants	130
Stimuli	18
Question-Asking Principal Component 2 (Variables):	
Stimulus-Gap Questions	0.01
Subjective-Gap Questions	0.64
Semantic Similarity: Questions-to-Stimuli	0.68
Question-Asking Principal Component 3 (Variables):	
Stimulus-Gap Questions	0.51
Subjective-Gap Questions	0.33
Semantic Similarity: Questions-to-Stimuli	–0.32

Note. The table depicts the results of the best fitting Linear Mixed-Effects model for state curiosity, with the question asking principal component 2, the question asking principal component 3, familiarity, and individual interest as fixed effects and participant and stimulus as random additive effects, expressed as Curiosity $\sim$ Question-Asking Principal Component 2 + Question-Asking Principal Component 3 + Familiarity + Individual Interest + (1 |Stimulus) + (1 |Participant). The model shows that question asking, familiarity, and individual interest positively and significantly predict curiosity. The standard error is displayed in parentheses under each coefficient. We also list the magnitude associated with each variable predictor that comprises the principal components (i.e., Stimulus-Gap Questions, Subjective-Gap Questions, and Semantic Similarity: Questions-to -Stimuli). ***p < 0.001; **p < 0.01; *p < 0.05

predicted by our second question-asking principal component ($\beta = 0.55$, 95% CI: 0.44–0.66, $SE = 0.06$, $p < .001$) and our third question-asking principal component ($\beta = 0.69$, 95% CI: 0.53–0.84, $SE = 0.08$, $p < .001$). The second question-asking principal component is comprised of the following variables and their associated magnitudes: Stimulus-Gap Questions (0.01), Subjective-Gap Questions (0.64), and Semantic Similarity: Questions-to-Stimuli (0.68). All of the variable predictors have the same sign as the regression coefficient, with higher magnitudes associated with Subjective-Gap Questions and Semantic Similarity:

Questions-to-Stimuli when compared to Stimulus-Gap Questions. This principal component emphasizes the relationship between curiosity and asking topic-related questions as well as questions that are semantically similar to the stimulus. The third question-asking principal component is comprised of the following variables and their associated magnitudes: Stimulus-Gap Questions (0.51), Subjective-Gap Questions (0.33), and Semantic Similarity: Questions-to-Stimuli (–0.32). The regression coefficient has the same sign as Stimulus-Gap Questions and Subjective-Gap Questions, but the opposite sign of Semantic Similarity: Questions-to-Stimuli. Therefore, this principal component emphasizes the relationship between curiosity and asking a high number of questions, that directly address gaps in the stimulus as well as topic-related questions. We also found that curiosity was predicted by familiarity ($\beta = 0.60$, 95% CI: 0.49–0.72, $SE = 0.06, p < .001$) and individual interest ($\beta = 0.42$, 95% CI: 0.31–0.54, $SE = 0.06, p < .001$). These results suggest that asking a high number of quality questions, having an individual interest in the topic, and being familiar with the information are associated with elevated levels of state curiosity.

Second, we analyzed our models of information foraging. We hypothesized that information foraging depends on both detecting gaps through question asking together with curiosity. For information foraging, we ran multiple models: a model for the number of times the participant looked at the stimulus-gap answers (Stimulus-Gap Foraging), a model for the number of times the participant looked at the answers to questions unrelated to the stimulus gaps (Subjective-Gap Foraging), and a model for whether the participant looked at the answer to each individual question (Single-Question Foraging). For Stimulus-Gap Foraging, we divided the coefficients by two because, for each stimulus, there were two buttons related to the gaps in the foraging phase; for Subjective-Gap Foraging, we divided the coefficients by 7 because seven buttons pertained to topic-related questions in the foraging phase.

Stimulus-Gap Foraging, as hypothesized, was predicted by both question asking and curiosity, with a BIC score of 4,946.02 (see Table 3). We found that Stimulus-Gap Foraging was significantly predicted by our first question-asking principal component ($\beta = 0.04$, 95% CI: 0.02–0.05, $SE = 0.01, p < .001$) and our third question-asking principal component ($\beta = 0.05$, 95% CI: 0.03–0.08, $SE = 0.01, p < .001$). The first question-asking principal component is comprised of the following variables and their associated magnitudes: Stimulus-Gap Questions (0.70), Subjective-Gap Questions (–0.39), Semantic Similarity: Questions-to-Stimuli (0.79), and Semantic Similarity: Grouped Questions-to-Foraging Questions (0.67). The regression coefficient has the same sign as Stimulus-Gap Questions, Semantic Similarity: Questions-to-Stimuli, and Semantic Similarity: Grouped Questions-to-Foraging Questions. This result suggests that Stimulus-Gap Foraging is promoted by asking Stimulus-Gap Questions that are semantically similar to both the stimulus and the foraging questions. The third question-asking principal component is comprised of the following variables and their associated magnitudes: Stimulus-Gap Questions (0.50), Subjective-Gap Questions (0.44), Semantic Similarity: Questions-to-Stimuli (0.005), and Semantic Similarity: Grouped Questions-to-Foraging Questions (–0.27). The regression coefficient has the same sign as Stimulus-Gap Questions, Subjective-Gap Questions, and Semantic Similarity: Questions-to-Stimuli, but the magnitude is much higher for Stimulus-Gap Questions and Subjective-Gap Questions. This principal component broadly indicates that Stimulus-Gap

Table 3
Regression coefficients predicting state information foraging from question asking and curiosity

	Stimulus-Gap Foraging	Subjective-Gap Foraging
(Intercept)	0.14***	0.40***
	(0.01)	(0.03)
Question-Asking Principal Component 1	0.04***	
	(0.01)	
Question-Asking Principal Component 2		0.05***
		(0.004)
Question-Asking Principal Component 3	0.05***	0.05***
	(0.01)	(0.006)
Residual Curiosity		0.06***
		(0.004)
Participant random effects	Y	Y
Stimulus random effects	Y	Y
AIC	4918.12	6861.36
BIC	4946.02	6900.41
Num. obs.	1958	1958
Participants	130	130
Stimuli	18	18
Question-Asking Principal Component 1 (Variables):		
Stimulus-Gap Questions	0.70	
Subjective-Gap Questions	–0.39	
Semantic Similarity: Questions-to-Stimuli	0.79	
Semantic Similarity: Grouped Quest-to-Foraging Quest	0.67	
Question-Asking Principal Component 2 (Variables):		
Stimulus-Gap Questions	–0.40	
Subjective-Gap Questions	0.78	
Semantic Similarity: Questions-to-Stimuli	0.31	
Semantic Similarity: Grouped Quest-to-Foraging Quest	0.52	
Question-Asking Principal Component 3 (Variables):		
Stimulus-Gap Questions	0.50	
Subjective-Gap Questions	0.44	
Semantic Similarity: Questions-to-Stimuli	0.005	
Semantic Similarity: Grouped Quest-to-Foraging Quest	–0.27	

Note. The table depicts the results of the best fitting Linear Mixed-Effects Model for information foraging, with question-asking principal component 1 (for Stimulus-Gap Foraging), question-asking principal component 2 (for Subjective-Gap Foraging), question-asking principal component 3 (for Stimulus-Gap and Subjective-Gap Foraging), and residual curiosity (for Subjective-Gap Foraging) as fixed effects and participant and stimulus as random additive effects. We ran models for multiple measures of information foraging. First, we ran a model for the number of stimulus-gap answers foraged for (i.e, Stimulus-Gap Foraging), expressed as Stimulus-Gap Foraging $\sim$ Question-Asking Principal Component 1 + Question-Asking Principal Component 3 + (1 |Stimulus) + (1 |Participant). Next, we ran a model for the number of forged-for answers that did not pertain to the

(*Continued*)

Table 3 (Continued)

gaps in the stimulus (i.e., Subjective-Gap Foraging), expressed as Subjective-Gap Foraging ∼ Question-Asking Principal Component 2 + Question-Asking Principal Component 3 + Residual Curiosity + (1 |Stimulus) + (1 |Participant). The standard error is reported under the coefficient. We also list the magnitude associated with each variable predictor that comprises the principal components (i.e., Stimulus-Gap Questions, Subjective-Gap Questions, Semantic Similarity: Questions-to-Stimuli, and Semantic Similarity: Grouped Questions-to-Foraging Questions). ***p < 0.001; **p < 0.01; *p < 0.05

Foraging is associated with asking a higher number of questions. Overall, these results suggest that Stimulus-Gap Foraging primarily depends on asking quality questions that are semantically relevant to the stimulus and to the questions participants could acquire answers to in the foraging phase. As expected, asking questions about the gaps in the stimulus more strongly predicted Stimulus-Gap Foraging than questions about the topic.

Subjective-Gap Foraging was predicted by asking questions and curiosity, with a BIC score of 6,900.41 (see Table 3). Subjective-Gap Foraging was significantly predicted by our second question-asking principal component ($\beta = 0.05, 95\%$ CI: 0.04–0.06, $SE = 0.004$, $p < .001$) and our third question-asking principal component ($\beta = 0.05, 95\%$ CI: 0.04–0.06, $SE = 0.006$, $p < .001$). The second question-asking principal component is comprised of the following variables and their associated magnitudes: Stimulus-Gap Questions (–0.40), Subjective-Gap Questions (0.78), Semantic Similarity: Questions-to-Stimuli (0.31), and Semantic Similarity: Grouped Questions-to-Foraging Questions (0.52). The regression coefficient is predominately influenced by Subjective-Gap Questions and Semantic Similarity: Grouped Questions-to-Foraging Questions, and to a lesser extent influenced by Semantic Similarity: Questions-to-Stimuli. These results suggest that Subjective-Gap Foraging is associated with asking Subjective-Gap Questions that are semantically related to the questions posed during the foraging phase. The third question-asking principal component is comprised of the following variables and their associated magnitudes: Stimulus-Gap Questions (0.50), Subjective-Gap Questions (0.44), Semantic Similarity: Questions-to-Stimuli (0.005), and Semantic Similarity: Grouped Questions-to-Foraging Questions (–0.27). We see that the regression coefficient has the same sign as Stimulus-Gap Questions, Subjective-Gap Questions, and Similarity: Questions-to-Stimuli. However, Similarity: Questions-to-Stimuli is much lower than the other variables in terms of magnitude. Generally speaking, this finding indicates the influence of asking a higher number of questions on Subjective-Gap Foraging. Subjective-Gap Foraging was also predicted by residual-state curiosity ($\beta = 0.06, 95\%$ CI: 0.05–0.06, $SE = 0.004$, $p < .001$). Subjective-Gap Foraging increased when the participants' typed questions were both relevant to the topic of the stimulus and semantically related to the stimulus and (especially) the foraging questions. As indicated by the results of the second question-asking principal component, Subjective-Gap Foraging was better predicted by the number of Subjective-Gap Questions asked than by the number of Stimulus-Gap Questions asked. These results suggest that when participants asked Subjective-Gap Questions that were later followed by the opportunity (in the foraging phase) to learn the answers to those

Table 4
Regression coefficients predicting memory from question asking, curiosity, and information interestingness

	Memory
(Intercept)	26.57***
	(3.50)
Question-Asking Principal Component 1	1.08***
	(0.32)
Question-Asking Principal Component 4	1.90**
	(0.60)
Residual Curiosity	1.45***
	(0.37)
Residual Information Interestingness	1.88***
	(0.37)
Participant random effects	Y
Stimulus random effects	Y
AIC	14,087.86
BIC	14,131.32
Num. obs.	1691
Participants	123
Stimuli	18
Question-Asking Principal Component 1 (Variables):	
Stimulus-Gap Questions	0.69
Subjective-Gap Questions	–0.35
Semantic Similarity: Questions-to-Stimuli	0.79
Semantic Similarity: Grouped Questions-to-Foraging Questions	0.69
Question-Asking Principal Component 4 (Variables):	
Stimulus-Gap Questions	0.21
Subjective-Gap Questions	0.05
Semantic Similarity: Questions-to-Stimuli	–0.47
Semantic Similarity: Grouped Questions-to-Foraging Questions	0.36

Note. The table depicts the results of the best fitting Linear Mixed-Effects Model for memory for gap answers, with question-asking principal component 1, question-asking principal component 4, residual curiosity, and residual information interestingness as fixed effects and participant and stimulus as random additive effects, expressed as Memory $\sim$ Question-Asking Principal Component 1 + Question-Asking Principal Component 4 + Residual Curiosity + Residual Information Interestingness + (1 |Stimulus) + (1 |Participant). The model shows that question asking, residual curiosity, and residual information interestingness significantly predict the participants' ability to recall information missing from the stimulus that they saw during the information foraging phase. The standard deviation is reported under the coefficient. We also list the magnitude associated with each variable predictor that comprises the principal components (i.e., Stimulus-Gap Questions, Subjective-Gap Questions, Semantic Similarity: Questions-to-Stimuli, and Semantic Similarity: Grouped Questions-to-Foraging Questions). ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$

questions, then they were more likely to seek out those answers. The likelihood of searching for answers also increased when the participants reported higher levels of state curiosity.

Single-Question Foraging was predicted by the semantic similarity between the participants' question and the foraging question, referred to as Semantic Similarity: Single

Question-to-Foraging Question ($\beta = 0.70$, 95% CI: 0.66–0.75, $SE = 0.02$, $p < .001$) as well as residual curiosity ($\beta = 0.26$, 95% CI: 0.23–0.30, $SE = 0.02$, $p < .001$), with a BIC score of 17,603.1. These results provide further evidence that participants are choosing to see answers to questions (in the foraging phase) that they themselves asked. In addition, we also see that curiosity continues to positively influence foraging behavior.

Third, we analyzed our memory model. We hypothesized that memory for gap-related information is influenced by initially asking questions about the missing information, curiosity to know more, and interest in or satisfaction with the gap-filling information (which we refer to as information interestingness). We wanted to only focus on cases where the participants saw the answers, so that their scores on the memory task reflected their exposure to the material as opposed to their prior knowledge. To do this, we removed instances where the participant did not look at either of the gap answers for a given stimulus (i.e., the information foraging score was equal to zero). Note that the average scores for instances where participants did not look at the gap answer was low (i.e., average score of 2.5 out of 100 points per stimulus). This also made it so that the information-interestingness rating was based (in part) on the answer of at least one of the gap questions. See Appendix S7 for means, standard deviations, and correlations.

Memory for gap answers was best predicted by question asking, curiosity, and information interestingness, with a BIC score of 14,131.32 (see Table 4). Specifically, memory was significantly predicted by our first question-asking principal component ($\beta = 1.08$, 95% CI: 0.45–1.70, $SE = 0.32$, $p < .001$) and our fourth question-asking principal component ($\beta = 1.90$, 95% CI: 0.71 –3.08, $SE = 0.60$, $p = .002$). The first question-asking principal component is comprised of the following variables and their associated magnitudes: Stimulus-Gap Questions (0.69), Subjective-Gap Questions (–0.35), Semantic Similarity: Questions-to-Stimuli (0.79), and Semantic Similarity: Grouped Questions-to-Foraging Questions (0.69). Stimulus-Gap Questions, Semantic Similarity: Questions-to-Stimulus, and Semantic Similarity: Grouped Questions-to-Foraging Questions have the same sign as the regression coefficient. This result indicates that memory is influenced by asking gap questions that relate to both the stimulus and foraging questions. The fourth question-asking principal component is comprised of the following variables and their associated magnitudes: Stimulus-Gap Questions (0.21), Subjective-Gap Questions (0.05), Semantic Similarity: Questions-to-Stimuli (–0.47), and Semantic Similarity: Grouped Questions-to-Foraging Questions (0.36). The regression coefficient shares the same sign with Stimulus-Gap Questions, Subjective-Gap Questions, and Semantic Similarity: Grouped Questions-to-Foraging Questions, with higher magnitudes associated with Stimulus-Gap Questions and Semantic Similarity: Grouped Questions-to-Foraging Questions. This suggests that memory is associated with asking questions about the gap that are semantically similar to the foraging questions. We also found that memory was significantly predicted by residual curiosity ($\beta = 1.45$, 95% CI: 0.73–2.16, $SE = 0.37$, $p < .001$) and residual information interestingness ($\beta = 1.88$, 95% CI: 1.16–2.60, $SE = 0.37$, $p < .001$). Overall, memory was predicted by the quantity and quality of questions asked, curiosity, and information interestingness. Specifically, memory for gap answers was better predicted by asking Stimulus-Gap Questions than asking Subjective-Gap Questions. This is to be expected because the memory task only asked about stimulus-gap

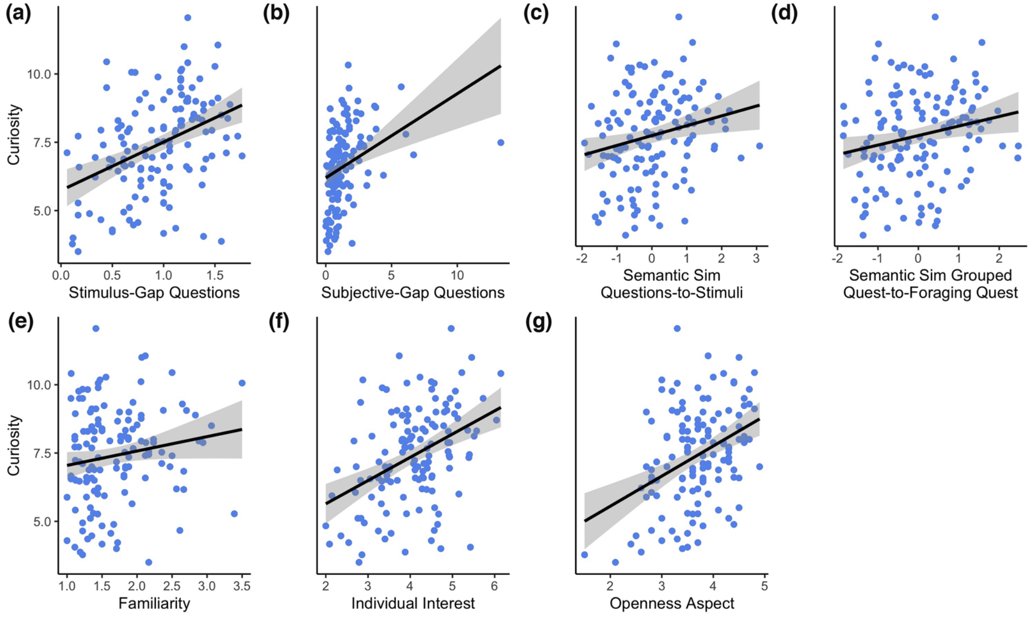

Fig. 4. Scatter plots for State Curiosity. Scatter plots suggest a positive relationship between State Curiosity and (a) Stimulus-Gap Questions, (b) Subjective-Gap Questions, (c) Semantic Similarity: Questions-to-Stimuli, (d) Semantic Similarity: Grouped Questions-to-Foraging Questions, (e) pre-experimental familiarity, (f) independently assessed individual interest, and (g) Openness Aspect from BFAS. Each point in the scatter plots represents the average score for each of the 135 participants. The solid line is the best fitting linear regression, and the shaded area is the 95% confidence interval.

answers. Also, participants were more likely to remember the information if they additionally judged the information they read (during the information foraging phase) to be interesting or satisfying. Surprisingly, state curiosity was not as strong of a predictor in the best fitting model, suggesting that it does not play as large a role in participants' ability to recall the missing information as does residual satisfaction with the acquired information.

3.2. *Impact of individual differences on learning*

For the individual difference analyses, our measures from the Curiosity Q&A Task were averaged and scaled between participants (see Figs. 4–7 for scatter plots). We ran linear regression models using the lm function (stats; R Core Team, 2013). We identified the best fitting models with the lowest BIC score using best subset selection, where we ran models for all combinations of predictors. If multiple models were not statistically significant (i.e., less than a 2-point difference in BIC score; Raftery, 1995), then we selected the model with the fewest parameters. After selecting the best model, we checked that each model met the assumptions for linear models (see Appendix S6 for details). See Table 5 for means, standard deviations, and correlations; and see Fig. 8 for a schematic overview of the results.

Table 5
Means, standard deviations, and correlations for individual difference analyses

	M	SD	1	2	3	4	5	6	7	8	9	10	11	12
1. Stimulus-Gap Questions	0.94	0.40	1.00											
2. Subjective-Gap Questions	1.36	1.56	0.53	1.00										
3. Semantic Similarity: Q-to-Stim	0.28	0.08	0.72	0.39	1.00									
4. Semantic Similarity: GQ-to-FQ	0.29	0.07	0.66	0.44	0.90	1.00								
5. Curiosity	7.41	1.77	0.40	0.34	0.22	0.21	1.00							
6. Familiarity	1.68	0.53	–0.11	0.18	–0.04	–0.08	0.16	1.00						
7. Individual Interest	4.08	0.85	0.10	0.13	0.12	0.05	0.41	0.12	1.00					
8. Openness Aspect	3.68	0.62	0.17	0.12	0.24	0.18	0.39	0.13	0.42	1.00				
9. Stimulus-Gap Foraging	1.26	0.44	0.72	0.47	0.55	0.55	0.44	–0.03	0.13	0.29	1.00			
10. Subjective-Gap Foraging	2.69	1.71	0.46	0.60	0.38	0.44	0.41	0.11	0.22	0.34	0.75	1.00		
11. Information Interestingness	4.19	0.89	0.17	0.24	0.11	0.11	0.78	0.23	0.41	0.46	0.39	0.45	1.00	
12. Memory	22.77	9.07	0.65	0.45	0.49	0.50	0.33	–0.01	0.14	0.25	0.82	0.60	0.33	1.00

Note. Variables used for the individual difference models of curiosity and information foraging. M and SD are used to represent mean and standard deviation, respectively where the Curiosity Q&A Task measures have been averaged per participant. The ranges for the measures are as follows: Stimulus-Gap Questions (0-2), Subjective-Gap Questions (there was no cap on the number of questions), Semantic Similarity: Questions-to-Stimuli (Q-to-Stim; 0-1), Semantic Similarity: Grouped Questions-to-Foraging Questions (GQ-to-FQ; 0-1), Curiosity (1-14)-combined measure of curiosity about questions (0-7) and curiosity about topic (1-7), Familiarity (1-7), Average Individual Interest (1-7), Stimulus-Gap Foraging (0-2), Subjective-Gap Foraging (0-7), Information Interestingness (1-7), and Memory (0-100).

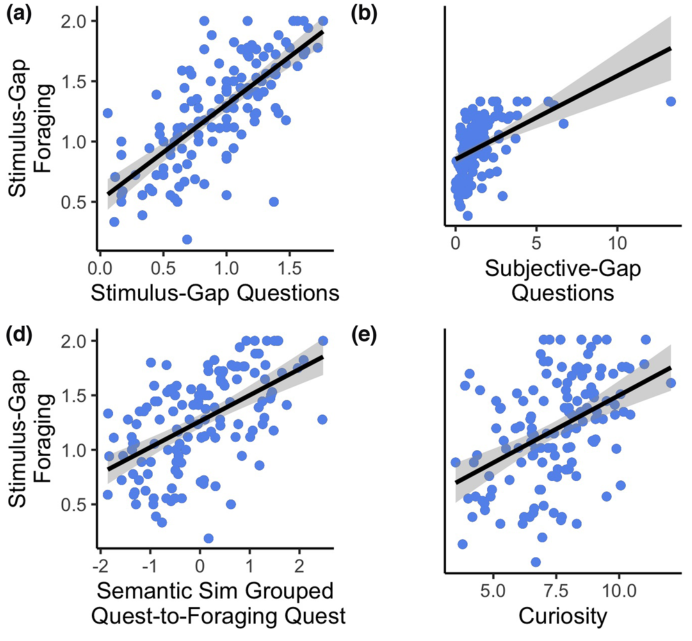

Fig. 5. Scatter plots for Stimulus-Gap Foraging. The plots suggest a positive relationship between Stimulus-Gap Foraging and (a) Stimulus-Gap Questions, (b) Subjective-Gap Questions, (c) Semantic Similarity: Questions-to-Stimuli, (d) Semantic Similarity: Grouped Questions-to-Foraging Questions, (e) State Curiosity, and (f) Openness Aspect from BFAS. Each point in the scatter plots represents the average score for each of the 135 participants. The solid line is the best fitting linear regression, and the shaded area is the 95% confidence interval.

First, we analyzed our model of state curiosity. We found that state curiosity was best predicted by question asking, individual interest, and openness to experience ($R^2 = .36$, Adjusted $R^2 = .34$; see Table 6). Curiosity is significantly predicted by our first question-asking principal component ($\beta = 0.36$, 95% CI: 0.18–0.53, $SE = 0.09$, $p < .001$) and our third question-asking principal component ($\beta = 0.77$, 95% CI: 0.28–1.26, $SE = 0.25$, $p = .002$). The first question-asking principal component is comprised of the following variables and their associated magnitudes: Stimulus-Gap Questions (0.90), Subjective-Gap Questions (0.72), and Semantic Similarity: Questions-to-Stimuli (0.85). All variables have the same sign and comparably strong magnitudes. The third question-asking principal component is comprised of the following variables and their associated magnitudes: Stimulus-Gap Questions (–0.40), Subjective-Gap Questions (0.11), and Semantic Similarity: Questions-to-Stimuli (0.34). Only Subjective-Gap Questions and Semantic Similarity: Questions-to-Stimulus had the same sign as the regression coefficient. We also found that curiosity was predicted by individual interest ($\beta = 0.51$, 95% CI: 0.24–0.78, $SE = 0.14$, $p < .001$) and Openness ($\beta = 0.39$, 95% CI: 0.12–0.67, $SE = 0.14$, $p = .006$). These results suggest that asking a high number of quality questions, having an individual interest in the topic, and the trait of openness are linked to increases in curiosity.

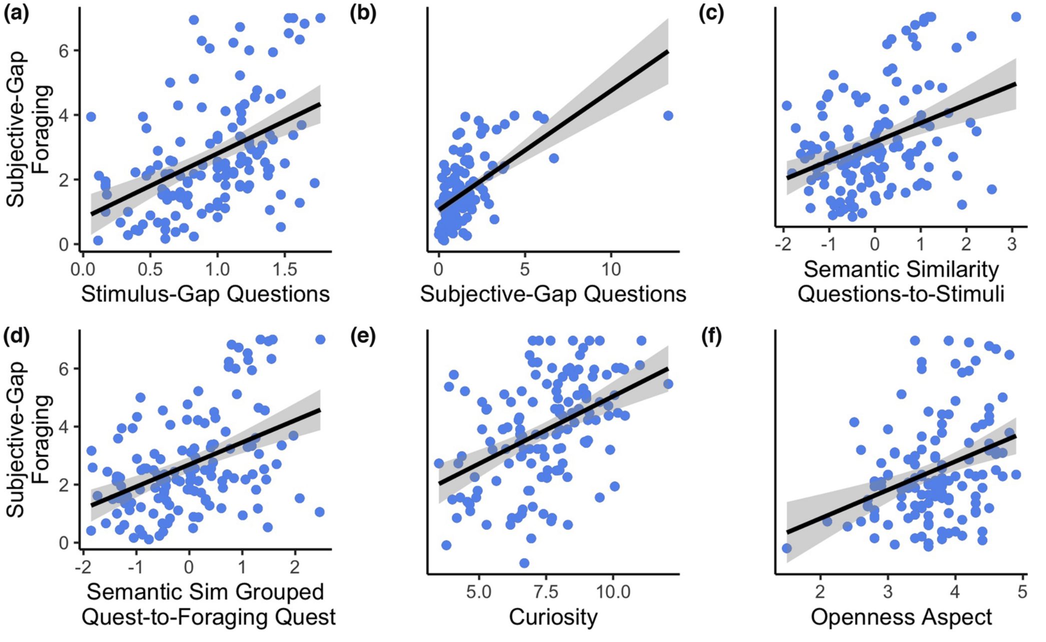

Fig. 6. Scatter plots for Subjective-Gap Foraging. The plots also suggest a positive relationship between Subjective-Gap Foraging and (a) Stimulus-Gap Questions, (b) Subjective-Gap Questions, (c) Semantic Similarity: Questions-to-Stimuli, (d) Semantic Similarity: Grouped Questions-to-Foraging Questions, (e) State Curiosity, and (f) Openness Aspect from BFAS. Each point in the scatter plots represents the average score for each of the 135 participants. The solid line is the best fitting linear regression, and the shaded area is the 95% confidence interval.

Next, we analyzed our models of information foraging. We ran two different linear models: one model that predicts foraging for information about the gaps in the stimulus (Stimulus-Gap Foraging) and one model that predicts foraging for information related to the stimulus topic (Subjective-Gap Foraging). For Stimulus-Gap Foraging, we divided the coefficients by two because, for each stimulus, there were two buttons related to the gaps in the foraging phase; for Subjective-Gap Foraging, we divided the coefficients by 7 because seven buttons pertained to topic-related questions in the foraging phase. For the Subjective-Gap Foraging model, we removed one influential point that was significantly influencing the model (Cook's Distance > 0.80; see Appendix S6 for more details).

Stimulus-Gap Foraging, as hypothesized, was predicted by both question asking and state curiosity ($R^2 = .55$, Adjusted $R^2 = .54$; see Table 7). Specifically, Stimulus-Gap Foraging was significantly predicted by our first question-asking principal component ($\beta = 0.09$, 95% CI: 0.07–0.10, $SE = 0.01$, $p < .001$) and our third question-asking principal component ($\beta = 0.09$, 95% CI : 0.05–0.14, $SE = 0.02$, $p < .001$). The first question-asking principal component is comprised of the following variables and their associated magnitudes: Stimulus-Gap Questions (0.87), Subjective-Gap Questions (0.66), Semantic Similarity: Questions-to-Stimuli (0.92), and Semantic Similarity: Grouped Questions-to-Foraging Questions (0.91). All of the variables are high in magnitude and share the same sign as the

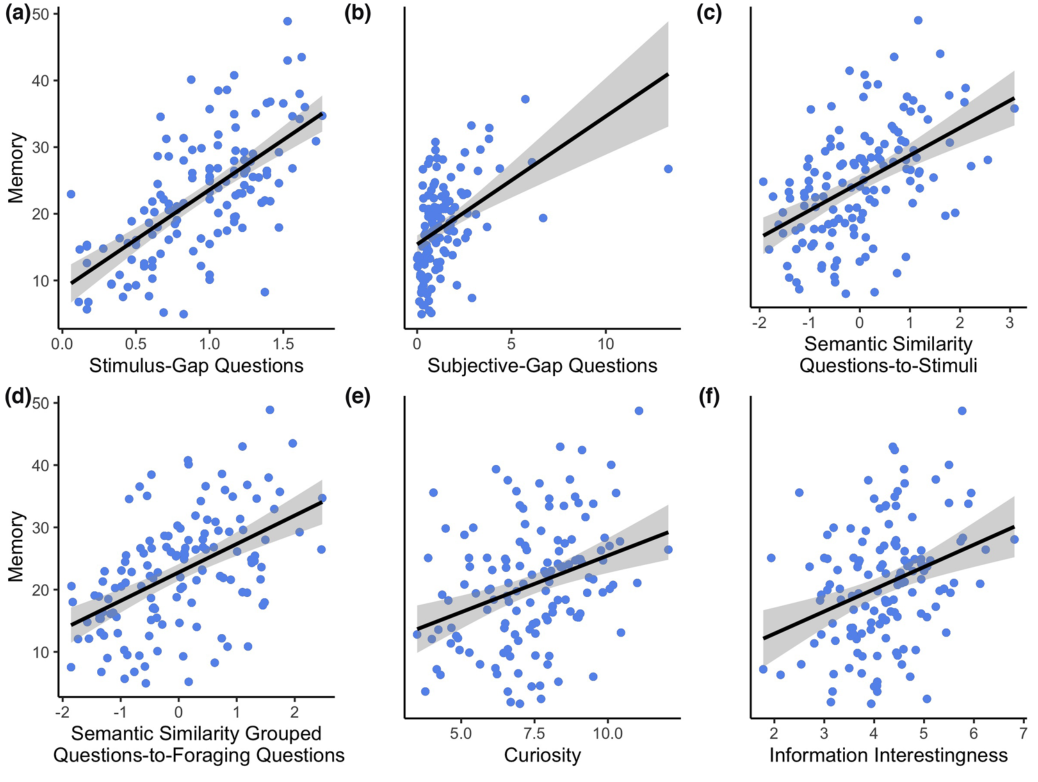

Fig. 7. Scatter plots for Memory Recall of Gap Information. Scatter plots suggest a positive relationship between Memory and (a) Stimulus-Gap Questions, (b) Subjective-Gap Questions, (c) Semantic Similarity: Questions-to-Stimuli, (d) Semantic Similarity: Grouped Questions-to-Foraging Questions, (e) State Curiosity, and (f) Information interestingness. Each point in the scatter plots represents the average score for each of the 135 participants. The solid line is the best fitting linear regression, and the shaded area is the 95% confidence interval.

regression coefficient, indicating that Stimulus-Gap Foraging is associated with asking a high number of quality questions that both relate to the stimulus and the foraging information. The third question-asking principal component is comprised of the following variables and their associated magnitudes: Stimulus-Gap Questions (0.49), Subjective-Gap Questions (−0.17), Semantic Similarity: Questions-to-Stimuli (−0.09), and Semantic Similarity: Grouped Questions-to-Foraging Questions (−0.26). Stimulus-Gap Questions is the only variable that has the same sign as the regression coefficient, indicating that Stimulus-Gap Foraging is influenced by asking Stimulus-Gap Questions. Stimulus-Gap Foraging was also predicted by residual curiosity after the influence of question asking had been removed ($\beta = 0.04, 95\%\ CI : 0.01 - 0.06, SE = 0.01, p = .003$). Overall, Stimulus-Gap Foraging was predicted by asking a high number of quality questions, especially those that pertained to the information missing from the stimulus, as well as curiosity to know more.

Subjective-Gap Foraging was predicted by question asking and openness to experience ($R^2 = .45$, Adjusted $R^2 = .44$; see Table 7). Specifically, we found that Subjective-Gap

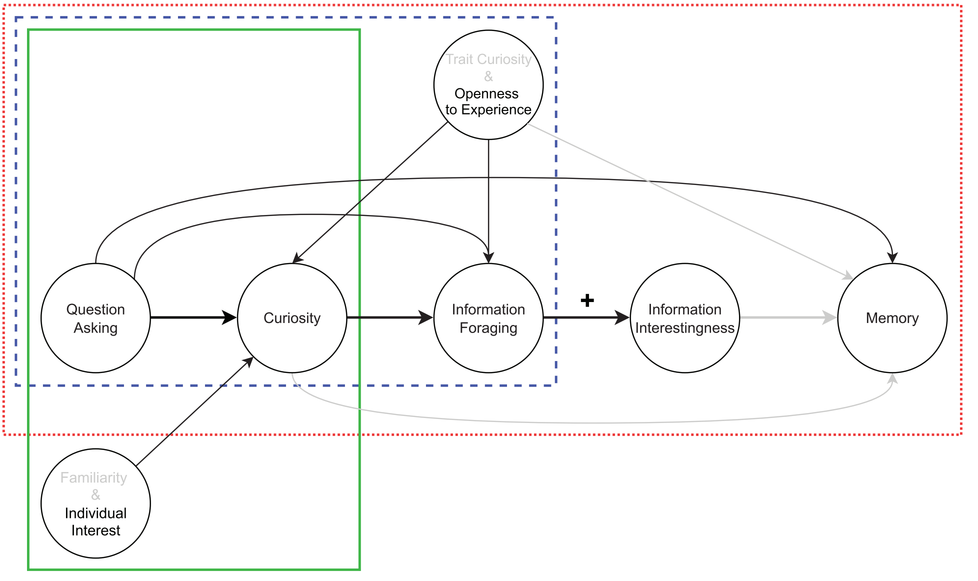

Fig. 8. Individual difference analyses of curiosity, information foraging, and memory. An overview of the active-curiosity-driven-learning model. Gray indicates that a predictor did not significantly predict the corresponding target variable(s).

Foraging was predicted by our first question-asking principal component ($\beta = 0.07$, 95% CI: 0.05–0.09, $SE = 0.01$, $p < .001$) and our second component ($\beta = 0.09$, 95% CI: 0.06–0.13, $SE = 0.02$, $p < .001$). As mentioned above, the first question asking principal component suggests that Subjective-Gap Foraging is associated with asking a high number of questions that are both semantically similar to the stimulus text and the questions presented during the foraging phase. The second question-asking principal component is comprised of the following variables and their associated magnitudes: Stimulus-Gap Questions (0.06), Subjective-Gap Questions (0.73), Semantic Similarity: Questions-to-Stimuli (–0.32), and Semantic Similarity: Grouped Questions-to-Foraging Questions (–0.27). We see that both Stimulus-Gap Questions and Subjective-Gap Questions share the same sign as the regression coefficient, and that Subjective-Gap Questions is much higher in magnitude than Stimulus-Gap Questions. Thus, this component emphasizes the strong relationship between asking Subjective-Gap Questions and selecting to see answers to Subjective-Gap Questions during the foraging phase. Subjective-Gap Foraging was also predicted by the openness aspect from BFAS ($\beta = 0.06$, 95% CI: 0.03–0.09, $SE = 0.02$, $p < .001$).

Finally, we analyzed our model of memory. We hypothesized that memory for gap-related information is influenced by initially asking questions about the missing information, curiosity to know more, and interest in or satisfaction with the gap-filling information (which we refer to as information interestingness). We wanted to only focus on cases where the participants saw the answers, so that their scores on the memory task reflected their exposure to the material as opposed to their prior knowledge. To do this, we removed instances where the

Table 6
Regression coefficients predicting individual differences in curiosity from question asking, familiarity, and individual interest

	Curiosity
(Intercept)	7.41***
	(0.12)
Question-Asking Principal Component 1	0.36***
	(0.09)
Question-Asking Principal Component 3	0.77**
	(0.25)
Individual Interest	0.51***
	(0.14)
Openness Aspect	0.39**
	(0.14)
R^2	0.36
Adj. R^2	0.34
Participants	135
Question-Asking Principal Component 1 (Variables):	
Stimulus-Gap Questions	0.90
Subjective-Gap Questions	0.72
Semantic Similarity: Questions-to-Stimuli	0.85
Question-Asking Principal Component 3 (Variables):	
Stimulus-Gap Questions	–0.40
Subjective-Gap Questions	0.11
Semantic Similarity: Questions-to-Stimuli	0.34

Note. The table depicts the results of the best fitting linear individual difference model for curiosity, expressed as Curiosity $\sim$ Question-Asking Principal Component 1 + Question-Asking Principal Component 3 + Individual Interest + Openness Aspect. The model shows that state curiosity was significantly predicted by question asking, individual interest, and the openness aspect from BFAS. The standard error is reported under the coefficient. We also list the magnitude associated with each variable predictor that comprises the principal components (i.e., Stimulus-Gap Questions, Subjective-Gap Questions, and Semantic Similarity: Questions-to-Stimuli). ***p < 0.001; **p < 0.01; *p < 0.05

participant did not look at either of the gap answers for a given stimulus (i.e., the information foraging score was equal to zero). This also made it so that the information interestingness rating was based (in part) on the answer of at least one of the gap questions. We identified the best fitting model using best subset selection. See Appendix S7 for means, standard deviations, and correlations.

For the individual difference model of memory, we found that memory for gap answers was predicted by question asking ($R^2 = .25$, Adjusted $R^2 = .24$; see Table 8). We found that memory was predicted by our first question-asking principal component ($\beta = 2.31$, 95% CI: 1.62–3.00, $SE = 0.35$, $p < .001$). The first question-asking principal component is comprised of the following variables and their associated magnitudes: Stimulus-Gap Questions

Table 7
Regression coefficients predicting individual differences in information foraging from question asking, curiosity, and openness

	Stimulus-Gap Foraging	Subjective-Gap Foraging
(Intercept)	0.63***	0.38***
	(0.01)	(0.02)
Question-Asking Principal Component 1	0.09***	0.07***
	(0.01)	(0.01)
Question-Asking Principal Component 2		0.09***
		(0.02)
Question-Asking Principal Component 3	0.09***	
	(0.02)	
Residual Curiosity	0.04**	
	(0.01)	
Openness Aspect		0.06***
		(0.02)
R^2	0.55	0.45
Adj. R^2	0.54	0.44
Participants	135	135
Question-Asking Principal Component 1 (Variables):		
Stimulus-Gap Questions	0.87	
Subjective-Gap Questions	0.66	
Semantic Similarity: Questions-to-Stimuli	0.92	
Semantic Similarity: Grouped Quest-to-Foraging Quest	0.91	
Question-Asking Principal Component 2 (Variables):		
Stimulus-Gap Questions	0.06	
Subjective-Gap Questions	0.73	
Semantic Similarity: Questions-to-Stimuli	–0.32	
Semantic Similarity: Grouped Quest-to-Foraging Quest	–0.27	
Question-Asking Principal Component 3 (Variables):		
Stimulus-Gap Questions	0.49	
Subjective-Gap Questions	–0.17	
Semantic Similarity: Questions-to-Stimuli	–0.09	
Semantic Similarity: Grouped Quest-to-Foraging Quest	–0.26	

Note. The table depicts the results of the linear regressions for the individual difference models of information foraging, with question asking and residual curiosity as fixed effects and participant and stimulus as random additive effects. We ran models for multiple measures of information foraging. First, we ran a model for the number of stimulus-gap answers foraged for (i.e, Stimulus-Gap Foraging), expressed as Stimulus-Gap Foraging $\sim$ Question-Asking Principal Component 1 + Question-Asking Principal Component 3 + Residual Curiosity. Next, we ran a model for the number of foraged-for answers that did not pertain to the gaps in the stimulus (i.e., Subjective-Gap Foraging), expressed as Subjective-Gap Foraging $\sim$ Question-Asking Principal Component 1 + Question-Asking Principal Component 2 + Openness Aspect. The model shows that Stimulus-Gap Foraging is significantly predicted by question asking and curiosity, whereas Subjective-Gap Foraging is significantly predicted by question asking and the openness aspect from BFAS. The standard error is reported under the coefficient. We also list the magnitude associated with each variable predictor that comprises the principal components (i.e., Stimulus-Gap Questions, Subjective-Gap Questions, Semantic Similarity: Questions-to-Stimuli, and Semantic Similarity: Grouped Questions-to-Foraging Questions). ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$

Table 8
Regression coefficients predicting individual differences in memory from question asking

	Memory
(Intercept)	26.01***
	(0.58)
Question-Asking Principal Component 1	2.31***
	(0.35)
R^2	0.25
Adj. R^2	0.24
Participants	135
Question-Asking Principal Component 1 (Variables):	
Stimulus-Gap Questions	0.86
Subjective-Gap Questions	0.64
Semantic Similarity: Questions-to-Stimuli	0.91
Semantic Similarity: Grouped Questions-to-Foraging Questions	0.92

Note. The table depicts the results of the best fitting linear model for memory for gap answers, with question asking as a fixed effect and participant and stimulus as random additive effects, expressed as Memory $\sim$ Question-Asking Principal Component 1. The model shows that question asking significantly predicts the participants' ability to recall information missing from the stimulus that they saw during the information foraging phase. The standard error is reported under the coefficient. We also list the magnitude associated with each variable predictor that comprises the principal component (i.e., Stimulus-Gap Questions, Subjective-Gap Questions, Semantic Similarity: Questions-to-Stimuli, and Semantic Similarity: Grouped Questions-to-Foraging Questions). ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$

(0.86), Subjective-Gap Questions (0.64), Semantic Similarity: Questions-to-Stimuli (0.91), and Semantic Similarity: Grouped Questions-to-Foraging Questions (0.92). For the first principal component, all question asking measures are positive in sign and high in magnitude. These results suggest that the individual differences in memory are dictated by the quality and quantity of question asking.

4. Discussion

Our goal was to understand how actively identifying gaps and posing questions to address those gaps influences curiosity and learning outcomes. Using our newly developed Curiosity Q&A paradigm, we found evidence supporting the broad role of question asking in setting the stage for all the subsequent steps of learning. Specifically, our results suggest that actively generating a high number of quality questions inspires curiosity, promotes information foraging, and enables subsequent memory. Our results also indicate that learning benefits from motivational states, such as curiosity to know more and satisfaction with the information acquired, as well as individual differences, such as openness to experience.

4.1. What sparks curiosity?

Curiosity is thought to be caused by both the recognition of an information gap and the subjective desire to obtain the missing information (Loewenstein, 1994). Here, we primarily examined whether identifying information gaps through question asking is associated with curiosity. Our findings suggest, as hypothesized, that generating questions is positively correlated with elevated curiosity. Using our newly developed Curiosity Q&A Task, we found that both the quantity of questions that were posed (asking a high number of questions that both target gaps in the stimulus and gaps in the participants' own knowledge) and the quality of those questions (semantic similarity between the questions asked and the stimuli) were linked to stronger levels of curiosity. These results explained fluctuations in curiosity for particular participants (measured by our state analyses) and individual differences in average curiosity across participants (measured by our individual difference analyses).

In line with information-gap theory (Loewenstein, 1994), these results emphasize the relationship between curiosity and the identification of information gaps. Our experimental design prohibits us from making a causal claim about the connection between question asking and curiosity. It is possible that the quality and quantity of questions asked causally influences curiosity. Support for this perspective comes from the findings of Golman et al. (2021), which suggests that curiosity arises from the desire to reduce uncertainty about the answer to a question. The strength of that desire depends on multiple factors, one of which is how relevant the question is to the context (referred to as salience). The questions asked by participants during the Curiosity Q&A Task are salient in multiple ways: (1) Subjective-Gap Questions are relevant to the context in that they target information missing from the stimulus; (2) both Stimulus-Gap and Subjective-Gap questions were asked in close proximity to when participants reported their curiosity; and (3) higher semantic similarity between the participants' typed questions and the stimulus text indicates that the questions are more in-line with the information from the stimulus. Therefore, the salience associated with asking a high number of quality questions may have directly influenced curiosity. Alternatively, it is possible that curiosity fueled the generation of a high number of quality questions. After reading the stimulus, participants may have initially experienced general curiosity about the topic, causing them to search the stimulus for clues about what they could learn next. Another possibility is a mixed causal relationship. At first, participants may pose a question that results in curiosity. Afterward, the resulting curiosity may cause them to search the stimulus or their prior knowledge for more gaps that can be translated into questions.

In our study, it is possible that participants who did not type questions failed to recognize that there were gaps in the stimulus or gaps in their prior knowledge relevant to the stimulus topic. If it is the case that question asking causes curiosity, their curiosity may be lower because they were unaware that there was more to know. Alternatively, participants who did not ask about the gaps may have sensed that something was missing from the statement, but were not able to articulate what it was (cf. Boyce-Jacino & DeDeo, 2020; McDaniel, Marsh, & Gouravajhala, 2022). In this case, the participants' self-reported curiosity may be lower because they were not exactly sure what information could be acquired. By clearly identifying information gaps through question asking, a person is able to appraise what information could be acquired, with the potential for curiosity to be sparked.

Our results also suggest that an individual's subjective relationship with the material (i.e., pre-experimental familiarity and independently assessed individual interest) impacts curiosity. Individual interest is a largely underexplored possible contributor to learning, where an individual's prior appreciation for a topic sparks approach behavior or state curiosity (Hidi & Renninger, 2006; Loewenstein, 1994). Recent work has sought to measure how interest in material translates to prolonged learning (Ainley et al., 2002). However, little is known about how general topics of interests (e.g., interest in biology) can impact curiosity. Hidi and Renninger's (2006) four-phase model of interest suggests that—across time and with increased experiential involvement—a person's interest transitions from transient situational amusement to an internalized aspirational orientation. That is, after individual interest in a topic has been established, an individual has the tendency to seek out and engage with information that pertains to the topic. Our results corroborate this claim, suggesting that individual interest motivates curiosity or desire to know more about a given topic.

Our results also indicate that curiosity can be positively influenced by an individual's pre-experimental familiarity with the information. We found that familiarity was a positive predictor of curiosity in our state regression model, but familiarity was not a predictor in our best fitting individual difference regression model. Familiarity may have been absent from our individual difference analyses because the entire range of values was not expressed at the aggregate level. Rather, participants were generally unfamiliar with the material presented in the task, leading to a low average score of familiarity ($mean = 1.68$ on a 7-point scale; $sd = 0.53$). Familiarity may have significantly influenced curiosity if our participants' average familiarity expressed the complete range of values, from low familiarity to high familiarity. Prior empirical and theoretical work commonly suggests that curiosity is highest when familiarity with, or uncertainty about, information is moderate (Baranes et al., 2015; Kang et al., 2009; Loewenstein, 1994). However, recent work suggests that high metacognitive familiarity or uncertainty predicts curiosity (Wade & Kidd, 2019). Our state regression model indicates a similar relationship to that of Wade and Kidd's results, where higher curiosity is associated with higher pre-experimental familiarity. But, our self-report familiarity rating did not explicitly address the gaps within the statement. Rather, participants were asked to generally report how familiar they were with the information before starting the experiment. Therefore, it is unclear whether this relationship reflects participants' perceived familiarity with the general topic of the statement or the gaps within the statement. If the positive impact of familiarity is based on the statement's general topic, this suggests that having prior-related knowledge about a given topic may amplify curiosity. Future research is needed to test whether familiarity with the missing information and familiarity with the general topic in which the gap arises both substantially contribute to curiosity.

In addition to question asking and individual interest, the trait of openness to experience also predicted curiosity in our individual difference analyses. Specifically, the openness aspect from BFAS (De Young et al., 2007). Openness to experience has been described as the tendency to welcome novelty and/or variety (Woo et al., 2013), and as "motivated cognitive flexibility" or "cognitive exploration" (DeYoung et al., 2005). This personality trait has also been associated with one's proclivity to engage with novel or varied intellectual information (e.g., concepts, facts, and puzzles) and to pursue new and diverse cultural, imaginative, and aesthetic experiences. In line with this characterization, our results suggest that openness to experience is associated with higher levels of curiosity or desire to know more information about a diverse set of topics.

Overall, these findings suggest that curiosity is positively correlated with both gap identification—through question asking—and an individual's subjective relationship with the material. Asking questions about the gaps indicates that a person has crafted a representation that accurately captures what is known and what is missing. The identification of an information gap facilitates the appraisal process, allowing for an individual to become curious. Familiarity and individual interest are two properties that support one's subjective desire to know more. That is, familiarity with or prior knowledge regarding the statement may make knowing the missing information more valuable, as either a way to further develop one's understanding of the topic or as a way to confirm that one's existing knowledge is correct. Individual interest, or prior engagement with the material indicates a person's appreciation for a topic, further promoting future engagement.

4.2. What supports information foraging?

Our findings, as hypothesized, suggest that asking questions about information gaps makes it more likely that a person will seek out the missing information. This was true for questions that were asked about the gaps in the stimulus (Stimulus-Gap Questions) and questions asked about gaps in the participants' prior knowledge (Subjective-Gap Questions). We ran multiple models for the foraging behaviors permitted by the task: the number of foraged-for answers that pertained to implicit information gaps within the stimulus (Stimulus-Gap Foraging), the number of foraged-for answers that had the potential to address the participants' Subjective-Gap Questions (Subjective-Gap Foraging), and whether the participant chose to look at the answer to each individual question in the foraging phase (Single-Question Foraging).

We found that Stimulus-Gap Foraging was better predicted by the number of Stimulus-Gap Questions asked, whereas Subjective-Gap Foraging was better predicted by the number of Subjective-Gap Questions asked. Although we did not directly map the Subjective-Gap Questions to questions in the foraging phase, as we did with the Stimulus-Gap Questions, the increased semantic similarity between the Subjective-Gap Questions and Subjective-Gap Foraging answers allows us to infer that participants were more likely to forage for subjective-gap answers when their questions matched the Subjective-Gap Foraging options. Our results also show that Stimulus-Gap Foraging is more highly correlated with Stimulus-Gap Questions (state analysis: $r = .23$; individual difference analysis: $r = .72$) than Subjective-Gap Questions (state analysis: $r = -.06$; individual difference analysis: $r = .46$). Similarly, Subjective-Gap Foraging is more highly correlated with Subjective-Gap Questions (state analysis: $r = .18$; individual difference analysis: $r = .60$) than Stimulus-Gap Questions (state analysis: $r = -.06$; individual difference analysis: $r = .47$). These findings indicate that the questions asked closely guided foraging behavior, wherein participants were more likely to seek relevant answers to questions they posed. This claim is further substantiated through our

Single-Question Foraging analysis, where we also found that participants were more likely to seek out answers to foraging questions that were semantically similar to the questions they typed. These results suggest that the higher the match was between the questions and the available answers to forage for, the more likely the participants were to search for answers. In line with research on self-directed learning (e.g., Gureckis & Markant, 2012; Kachergis, Rhodes, & Gureckis, 2017), question asking directed the participants' attention to gap-related information when they had the opportunity to forage.

Our results are pertinent to a recent debate in the curiosity literature regarding how best to interpret an information gap. On the one hand, some researchers claim that curiosity is driven by the potential information gain associated with collecting additional information, and do not frame the missing information in the form of an answer to a question (Gottlieb et al., 2013; Van de Cruys et al., 2021). On the other hand, recent work posits that curiosity is motivated by the desire to find out answers to unanswered questions (Golman & Loewenstein, 2018; Golman et al., 2021), which have the potential to help us make sense of the world (Chater & Loewenstein, 2016; Liquin & Lombrozo, 2020). Our results support the claim that information gaps can be interpreted as questions for which the answer is unknown or uncertain, causing people to venture forth in search of the answer.

We also saw that curiosity—both about the questions participants asked (with the potential of some of those questions pertaining to the gaps in the stimulus) and curiosity to know more about the stimulus—promotes foraging for the missing information. Specifically, curiosity was a significant predictor of Stimulus-Gap Foraging (for our individual difference analyses), Subjective-Gap Foraging (for our state analysis), and Single-Question Foraging. These results corroborate studies showing that people choose to wait longer for or seek out information that they are curious to know more about (Baranes et al., 2015; Litman et al., 2005; Marvin & Shohamy, 2016), and are consistent with a view of curiosity as a process that proactively motivates individuals to seek new information.

Our state-level model for Stimulus-Gap Foraging and our individual difference model for Subjective-Gap Foraging did not have curiosity as a significant predictor. Instead, our individual difference model of Subjective-Gap Foraging was predicted by the trait of openness to experience, specifically the openness aspect from BFAS (De Young et al., 2007). Being high in openness may promote the acquisition of diverse information that is less tethered to the task or gaps implicitly missing from the environment. Unlike the Stimulus-Gap Foraging—where sought-after information is directly connected to gaps in the stimulus—Subjective-Gap Foraging involves seeking out information that is relevant to the topic but not as directly rooted in the meaning of the stimulus or task. Our results suggest that openness to experience can manifest in exploring divergent information that is loosely related to the current context.

Overall, these results highlight the significant role that both question asking and curiosity play during information foraging. Asking a higher number of semantically relevant questions indicates that you have spent more time and more precisely or accurately represented the current situation, and were able to use the representation to detect what was missing and self-direct learning pursuits. Additionally, curiosity motivates information foraging by signaling whether the gaps are viewed as important or sufficiently meaningful to attempt to ameliorate. Together, foraging results as a conjunction of detecting gaps and finding them valuable enough to want to fill them.

4.3. *What causes superior memory?*

We found support for our hypothesis that question asking is associated with superior memory for gap answers. The quality and quantity of questions play a predominant role in our state and individual analyses. Therefore, in line with the generation effect (e.g., Bertsch et al., 2007; Rosner et al., 2013; Slamecka & Graf, 1978) and evidence on the learning and memory benefits of "desirable difficulty" (e.g., Kachergis et al., 2017), one could infer that self-guided inquiry leads to heightened recall of information. Our results also extend the benefits of self-generating a question on memory as seen when used as a study technique (Foos et al., 1994) and when generating conceptual level questions (Bugg & McDaniel, 2021). Because the information gaps in the Curiosity Q&A Task were implicit in the stimuli, it is likely that the participants needed to create a high-quality representation to identify what was missing. We speculate that the effort put into parsing the information and building the representation (perhaps through a process of covert self-explanation, cf. Ruggeri, Xu, & Lombrozo, 2019) was the primary difference between participants who did and did not ask the gap questions. Furthermore, initially representing the information gaps, upon reading the stimulus, may have made it easier to integrate the missing information, once acquired. Therefore, the quality of the initial representation may have supported the integration and encoding of the missing information. We find indirect support for the beneficial effect of high-quality representations on memory through our measures of the quality of questions asked. High-quality questions—that are semantically similar to the stimulus and foraged-for information—imply that the participant has generated a high-quality representation, allowing them to effectively capture incoming information in question form. Therefore, the association between high-quality questions and later successful memory recall suggests that memory may also be associated with high-quality representations.

For our state-level analyses, we find that both curiosity and residual interest or satisfaction with the acquired information increases later memory recall. Curiosity has previously been shown to enhance subsequent memory for trivia facts (Gruber et al., 2014; Kang et al., 2009; Ligneul et al., 2018) and for incidental information presented concurrently with trivia questions (Gruber et al., 2014). However, our state regression model suggests that the impact of curiosity on memory may not be as strong as residual information interestingness. Our residual information interestingness measure results from removing the influence of curiosity from the reported interest or satisfaction with the acquired information. The new measure computes the prediction error, wherein the remaining satisfaction is not captured by the anticipated satisfaction or curiosity. Our results suggest that, rather than the largely future-directed anticipated satisfaction of learning something (i.e., curiosity), the prediction error in satisfaction with the information more strongly predicts subsequent memory recall. McGillivray, Murayama, and Castel (2015) reached a somewhat similar conclusion, finding that participants' memory was best predicted by postanswer interest in the trivia answers, instead of curiosity. Similar to our findings, Marvin and Shohamy (2016) found that the prediction error, calculated by subtracting answer satisfaction (or information interestingness) from initial curiosity, predicted memory of trivia answers. That is, when answer satisfaction, or interest, was higher than initial curiosity, the participants were more likely to recall the information. In line with our findings, this result suggest that a prediction error in one's interest in the acquired information predicts subsequent memory, beyond the initial curiosity to know the information.

4.4. Limitations

Despite its strengths, including the provision of multiple vantage points on how question asking shapes learning, our experimental design has certain limitations. One limitation is that we did not have a self-reported state curiosity measure for every question that participants asked. Instead, our measure asked participants to evaluate their curiosity level regarding the subset of questions they had generated for a given factual statement, and also to rate their curiosity about the topic of each factual statement. Eliciting self-reported curiosity ratings for each specific question that a participant generated would allow for a direct measure of the role that curiosity played when the question was asked. Nonetheless, our question and topic-based measure of curiosity was still tied to each of the factual stimuli, and enabled us to also account for the level of curiosity even when a participant did not ask a question. Similarly, we did not ask participants to assess how interesting or satisfying they found each specific gap-related answer seen during the information-foraging phase, but rather how interesting they found the subset of information they had foraged for in relation to each factual stimulus. Eliciting such "answer specific" interestingness ratings would have more accurately accounted for how particular interest in an answer related to subsequent memory. Still, our subset measure successfully captured a strong positive relationship between information interestingness and memory, which we speculate would intensify if we had an even more tightly yoked measure of interestingness for each gap answer. A third limitation is that the presentation of the answer was slightly postponed until participants had generated all of their questions for a given factual stimulus. It is unknown how these brief delays in satisfying their curiosity may have augmented, or diminished, the intensity of participants' curiosity, or influenced the specific pattern of their subsequent information-foraging behaviors. A fourth limitation is the number and diversity of stimulus provided. Given that we wanted to collect quality data within a reasonable amount of time, we developed 18 stimuli on a diverse set of topics. Although we attempted to include a range of topics that related to a variety of individual interests, it is possible that some of the participants' interests were not sufficiently captured by the set of stimuli. Nonetheless, the 18 stimuli were, in fact, sufficiently varied and broad ranging to capture the interest of the vast majority of participants. For example, out of 135 participants, the participants on average asked at minimum one question in response to at least 11 of the stimuli, and chose to see the answer to (at minimum) one foraging question in response to at least 13 of the stimuli.

4.5. Concluding thoughts and future research

Asking the "right" question allows for scientific advancements and facilitates learning (Coenen et al., 2019). Accordingly, our research suggests that posing questions scaffolds knowledge acquisition and improves subsequent memory. But, what leads us to become aware that there are other, not previously recognized, new or emergent gaps that we now also should be cognizant of as we move forward in our explorations or endeavors? Question asking is the product of underexplored processes, where one generates a representation that allows for gap detection. By shifting our viewpoint to focus on the mental representations in which gaps arise, we can improve our understanding of the cognitive processes and neural substrates that facilitate comprehension and learning (McDaniel et al., 2022; Tawfik et al., 2020). Future research can provide more direct measures of an individual's mental representation, allowing for the formalization of information gaps (e.g., missing entities and relationships).

In conclusion, it is clear that curiosity drives us to pursue missing information. But, how is our pursuit impacted by whether the information sought provides answers to our own questions versus questions posed by other people? And how much do our subjective feelings of not only curiosity, but also our interest in the revealed information, shape what we learn? To answer these questions, we tested the influence of actively generating a question in a learning context using our newly developed Curiosity Q&A Task that combines perspectives from both self-generated and curiosity-driven learning. We found that participants who generated a high number of quality questions about information gaps expressed higher levels of curiosity, were more likely to forage for the missing information, and were more likely to remember the information they acquired. We also found that (1) curiosity was further enhanced by pre-experimental familiarity and individual interest; (2) curiosity provided additional motivation for information foraging; and (3) that both curiosity and satisfaction with the acquired information boosted memory recall. But, curiosity was not simply a mediator of the down-stream effects of question asking on learning. Indeed, question asking emerged as a more efficient predictor of memory than (topic and question-based) curiosity. This is an important piece of indirect evidence supporting the key role of question asking in setting the stage for all the subsequent steps of active-curiosity-driven learning. Asking questions enhances the value of missing information and renders a mental representation that is easier to modify and extend once the new information is acquired, with important implications for learning and discovery of all forms.

Acknowledgments

The authors thank Joshua Gangl, Eunjeong Lee, John Martin, John McDermid, Sandy Nguyen, Pauline Ojambo, Mehvish Safdar, Roberto Strabala, Carrie Thomas, Grace Vieth, Nguyen Vu, and the Thinking Lab members. Portions of the results from the pilot study were reported, in poster format, at the 2018 annual meeting of the American Association for the Advancement of Science in Austin, TX and at the 2018 annual meeting of the Association for Psychological Science in San Francisco, CA; and as a flash talk at the Exploring Curiosity Conference, Nov. 22–23, 2018, in Amsterdam, Netherlands.

References

Ainley, M., Hidi, S., & Berndorff, D. (2002). Interest, learning, and the psychological processes that mediate their relationship. *Journal of Educational Psychology*, *94*(3), 545–561.

Baranes, A., Oudeyer, P.-Y., & Gottlieb, J. (2015). Eye movements reveal epistemic curiosity in human observers. *Vision Research*, *117*, 81–90.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Berlyne, D. E. (1960). *Conflict, arousal, and curiosity*. McGraw-Hill Book Company.

Bertsch, S., Pesta, B. J., Wiscott, R., & McDaniel, M. A. (2007). The generation effect: A meta-analytic review. *Memory & Cognition*, *35*, 201–210.

Boyce-Jacino, C., & DeDeo, S. (2020). Opacity, obscurity, and the geometry of question-asking. *Cognition*, *196*, 104071.

Bugg, J. M., & McDaniel, M. A. (2021). Selective benefits of question self-generation and answering for remembering expository text. *Journal of Educational Psychology*, *104*, 922–931.

Chater, N., & Loewenstein, G. (2016). The under-appreciated drive for sense-making. *Journal of Economic Behavior & Organization*, *126*(Part B), 137–154.

Coenen, A., Nelson, J. D., & Gureckis, T. M. (2019). Asking the right questions about the psychology of human inquiry: Nine open challenges. *Psychonomic Bulletin and Review*, *26*, 1548–1587.

De Young, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the big five. *Journal of Personality and Social Psychology*, *93*(5), 880–896.

DeYoung, C. G., Peterson, J., & Higgins, D. (2005). Sources of openness/intellect: Cognitive and neuropsychological correlates of the fifth factor of personality. *Journal of Personality*, *73*, 825–858.

Dillon, J. T. (1988). The remedial status of student questioning. *Journal of Curriculum Studies*, *20*(3), 197–210.

Foos, P. W., Mora, J. J., & Tkacz, S. (1994). Student study techniques and the generation effect. *Journal of Educational Psychology*, *88*, 567–576.

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression (3rd ed.)*. Thousand Oaks, CA: Sage.

Golman, R., & Loewenstein, G. (2018). Information gaps: A theory of preferences regarding the presence and absence of information. *Decision*, *5*(3), 143–164.

Golman, R., Loewenstein, G., Molnar, A., & Saccardo, S. (2021). The demand for, and avoidance of, information. *Management Science*, *68*(9), 6454–6476.

Gottlieb, J., Oudeyer, P. Y., Lopes, M., & Baranes, A. (2013). Information-seeking, curiosity, and attention: Computational and neural mechanisms. *Trends in Cognitive Sciences*, *17*(11), 585–593.

Graesser, A. C., & Olde, B. A. (2003). How does one know whether a person understands a device? The quality of questions the person asks when the device breaks down. *Journal of Educational Psychology*, *95*, 524–536.

Graesser, A. C., & Person, N. K. (1994). Question asking during tutoring. *American Educational Research Journal*, *31*(1), 104–137.

Gruber, M. J., Gelman, B. D., & Ranganath, C. (2014). States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron*, *84*, 486–496.

Gruber, M. J., & Ranganath, C. (2019). How curiosity enhances hippocampus-dependent memory: The prediction, appraisal, curiosity, and exploration (PACE) framework. *Trends in Cognitive Sciences*, *23*(12), 1014–1025.

Gureckis, T. M., & Markant, D. B. (2012). Self-directed learning: A cognitive and computational perspective. *Perspectives on Psychological Science*, *7*, 464–481.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, *41*(2), 111–127.

Jirout, J. J. (2020). Supporting early scientific thinking through curiosity. *Frontiers in Psychology*, *11*, 1–7. https://www.frontiersin.org/articles/10.3389/fpsyg.2020.01717/full

Johnson-Laird, P. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cognitive science series. Harvard University Press.

Kachergis, G., Rhodes, M., & Gureckis, T. (2017). Desirable difficulties during the development of active inquiry skills. *Cognition*, *166*, 407–417.

Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T., & Camerer, C. F. (2009). The wick in the candle of learning: Epistemic curiosity activates reward circuitry and enhances memory. *Psychological Science*, *20*, 963–973.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*, 163–182.

Ligneul, R., Mermillod, M., & Morisseau, T. (2018). From relief to surprise: Dual control of epistemic curiosity in the human brain. *NeuroImage*, *181*, 490–500.

Liquin, E. G., & Lombrozo, T. (2020). A functional approach to explanation-seeking curiosity. *Cognitive Psychology*, *119*, 101276.

Litman, J. (2008). Interest and deprivation factors of epistemic curiosity. *Personality and Individual Differences*, *44*, 1585–1595.

Litman, J., Hutchins, T., & Russon, R. (2005). Epistemic curiosity, feeling-of-knowing, and exploratory behaviour. *Cognition and Emotion*, *19*(4), 559–582.

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, *116*, 75–98.

Lydon-Staley, D., Zhou, D., Blevins, A. S., Zurn, P., & Bassett, D. (2021). Hunters, busybodies and the knowledge network building associated with deprivation curiosity. *Nature Human Behavior*, *5*, 327–336.

Markant, D. B. (2020). Active transitive inference: When learner control facilitates integrative encoding. *Cognition*, *200*, 104188.

Markant, D. B., DuBrow, S., Davachi, L., & Gureckis, T. M. (2014). Deconstructing the effect of self-directed study on episodic memory. *Memory and Cognition*, *42*, 1211–1224.

Markant, D. B., Ruggeri, A., Gureckis, T. M., & Xu, F. (2016). Enhanced memory as a common effect of active learning. *Mind, Brain, and Education*, *10*(3), 142–152.

Marvin, C. B., & Shohamy, D. (2016). Curiosity and reward: Valence predicts choice and information prediction errors enhance learning. *Journal of Experimental Psychology: General*, *145*, 266–272.

McDaniel, M. A., Marsh, E. J., & Gouravajhala, R. (2022). Individual differences in structure building: Impacts on comprehension and learning, theoretical underpinnings, and support for less able structure builders. *Perspectives on Psychological Science*, *17*, 385–406.

McGillivray, S., Murayama, K., & Castel, A. D. (2015). Thirst for knowledge: The effects of curiosity and interest on memory in younger and older adults. *Psychology and Aging*, *30*, 835–841.

Philip, S., Shola, P., & John, A., O. (2014). Application of content-based approach in research paper recommendation system for a digital library. *International Journal of Advanced Computer Science and Applications*, *5*, 37–40.

Pirolli, P. (2005). Rational analyses of information foraging on the web. *Cognitive Science*, *29*, 343–373.

Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review*, *106*, 643–675.

Psychology Software Tools, I. (2016). E-prime. https://support.pstnet.com/

R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–163.

Rosner, Z. A., Elman, J. A., & Shimamura, A. P. (2013). The generation effect: Activating broad neural circuits during memory encoding. *Cortex*, *49*, 1901–1909.

RStudio Team. (2020). *RStudio: Integrated Development Environment for R*. Boston, MA: RStudio, PBC.

Ruggeri, A., Xu, F., & Lombrozo, T. (2019). Effects of explanation on children's question asking. *Cognition*, *191*, 103966. https://doi.org/10.1016/j.cognition.2019.05.003

Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, *24*(5), 513–523.

Savolainen, R. (2018). Berrypicking and information foraging: Comparison of two theoretical frameworks for studying exploratory search. *Journal of Information Science*, *44*(5), 580–593.

Sharot, T., & Sunstein, C. R. (2020). How people decide what they want to know. *Nature Human Behaviour*, *4*, 14–19.

Shibata, N., Kajikawa, Y., & Sakata, I. (2012). Link prediction in citation networks. *Journal of the American Society for Information Science and Technology*, *63*(1), 78–85.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, *4*, 592–604.

Tawfik, A., Graesser, A., Gatewood, J., & Gishbaugher, J. (2020). Role of questions in inquiry-based instruction: Towards a design taxonomy for question-asking and implications for design. *Educational Technology Research and Development*, *68*, 1–25.

Van de Cruys, S., Damiano, C., Boddez, Y., Król, M., Goetschalckx, L., & Wagemans, J. (2021). Visual affects: Linking curiosity, Aha-Erlebnis, and memory through information gain. *Cognition*, *212*, 104698.

Wade, S., & Kidd, C. (2019). The role of prior knowledge and curiosity in learning. *Psychonomic Bulletin & Review*, *26*, 1377–1387.

Woo, S. E., Chernyshenko, O., Longley, A., Zhang, Z.-X., Chiu, C. Y., & Stark, S. (2013). Openness to experience: Its lower level structure, measurement, and cross-cultural equivalence. *Journal of Personality Assessment*, *96*, 29–45.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Supporting Information