Cognitive Psychology 119 (2020) 101276

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

A functional approach to explanation-seeking curiosity

Emily G. Liquin*, Tania Lombrozo

Department of Psychology, Princeton University, Peretsman Scully Hall, Princeton, NJ 08540, USA

A R T I C L E I N F O

Keywords:
Explanation
Curiosity
Exploration
Causal attribution
Question-asking

A B S T R A C T

Why do some (and only some) observations prompt people to ask "why?" We propose a functional approach to "Explanation-Seeking Curiosity" (ESC): the state that motivates people to seek an explanation. If ESC tends to prompt explanation search when doing so is likely to be beneficial, we can use prior work on the functional consequences of explanation search to derive "forward-looking" candidate triggers of ESC—those that concern expectations about the downstream consequences of pursuing explanation search. Across three studies (N = 867), we test hypotheses derived from this functional approach. In Studies 1–3, we find that ESC is most strongly predicted by expectations about future learning and future utility. We also find that judgments of novelty, surprise, and information gap predict ESC, consistent with prior work on curiosity; however, the role for forward-looking considerations is not reducible to these factors. In Studies 2–3, we find that predictors of ESC form three clusters, expectations about learning (about the target of explanation), expectations about export (to other cases and future contexts), and backward-looking considerations (having to do with the relationship between the target of explanation and prior knowledge). Additionally, these clusters are consistent across stimulus sets that probe ESC, but not fact-seeking curiosity. These findings suggest that explanation-seeking curiosity is aroused in a systematic way, and that people are not only sensitive to the match or mismatch between a given stimulus and their current or former beliefs, but to how they expect an explanation for that stimulus to improve their epistemic state.

1. Introduction

Why do humans ask why? Much of human learning, from childhood through adulthood, is achieved by asking questions. Children begin to ask information-seeking questions before the age of two, and they ask increasingly more *explanation*-seeking questions (often "why" questions) between the ages of two and three (Chouinard, 2007; Hickling & Wellman, 2001). Explanation remains an important mechanism for learning throughout the human lifespan: children ages two to six use self-generated explanations to guide their exploration (Legare, 2012), and adults use explanatory principles to guide their inferences (Lombrozo, 2016).

Despite the prevalence of "why" questions and the importance of explanation, humans do not always ask questions when they lack an explanation. In fact, there are many things in the world we *do not* question: why a particular tree is 57 feet tall, why dogs have eyes, or why a certain conference room has an odd number of chairs (to offer only a small number of possible cases). These examples illustrate that our *explanation-seeking curiosity* (ESC) is highly selective—humans are "why-curious" about some things, but not about others. Of course, there may be some circumstances under which these observations *would* prompt explanation-seeking: for example, if a 57-foot tree were surrounded by trees that were 56 feet tall, or for a person who expects trees of that species not to exceed 50 feet

* Corresponding author.
E-mail addresses: eliquin@princeton.edu (E.G. Liquin), lombrozo@princeton.edu (T. Lombrozo).

https://doi.org/10.1016/j.cogpsych.2020.101276
Received 4 August 2019; Received in revised form 23 December 2019; Accepted 20 January 2020
Available online 12 February 2020
0010-0285/ © 2020 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/BY-NC-ND/4.0/).

in height.[1] In other words, ESC is highly context-dependent and person-dependent, as well. However, little empirical work has addressed what distinguishes those things that elicit why-curiosity from those that do not, or what explains why ESC differs across contexts and individuals.

In the current paper, we develop and test a novel set of predictions about what triggers explanation-seeking curiosity. We call our approach a "functional" approach because our predictions derive from considering the functional consequences of explanation-seeking within the process of inquiry. Within this process, observations or facts will sometimes spark a learner's explanation-seeking curiosity, and thereby motivate explanation search, whether it take the form of reflection, exploration, or asking others. Importantly, explanation-seeking has a distinctive profile of functional consequences: as we argue below, explanation-seeking facilitates learning about the explanandum (i.e., what is being explained), learning for export (i.e., to generalize beyond the explanandum), and learning from the domain-based expertise of others. If ESC is in part sensitive to how likely it is that these functional consequences will be achieved, then *expectations* about these consequences—whether implicit or explicit—could guide when ESC is experienced and to what degree. In other words, learners should experience (more) ESC for observations or facts that, if explained, would likely support learning about the explanandum, for export, and/or from the domain-based expertise of others.

We begin, in the section that follows, by noting the most relevant previous research on information search and curiosity. We then turn to previous literature on the functional consequences of explanation search, motivating the three functions indicated above. For each function, we derive corresponding triggers of ESC, many of which have not been posited or investigated in prior research. We next evaluate these new triggers, along with potential triggers from prior research, by testing how well they predict ESC across diverse sets of questions. Specifically, we report three experiments employing a novel method that combines the diversity and ecological validity of real-world stimuli with the rigors of the lab: we sample explanation-seeking questions posed by users on the internet, and we have new groups of participants indicate their curiosity about the answers, evaluate the questions for the presence of candidate triggers, and/or reveal a subset of answers. In this way we can identify which triggers predict ESC. Finally, we explore (in Study 3) how well these potential triggers predict *fact-seeking curiosity* (FSC) in comparison to ESC, to help define the contours of our approach.

1.1. Prior approaches to curiosity and information-seeking

Mirroring our functional approach, a subset of prior work on question-asking and curiosity has focused on learners' expectations about the consequences of seeking information. We characterize such considerations as "forward-looking," as they involve an evaluation of the *expected outcomes* of pursuing information, as opposed to "backward-looking" considerations (such as novelty or surprise) which involve an evaluation of how a stimulus fits in with a learner's prior beliefs. For example, research on question-asking has considered how questions are shaped by their potential to solve problems (Chouinard, 2007; Legare, Mills, Souza, Plummer, & Yasskin, 2013) and reduce uncertainty (Rothe, Lake, & Gureckis, 2018; Ruggeri & Lombrozo, 2015; Ruggeri, Lombrozo, Griffiths, & Xu, 2016; Ruggeri, Sim, & Xu, 2017).[2] Turning to curiosity, the influential *information gap* theory (Loewenstein, 1994) is also articulated as forward-looking, focusing on the discrepancy between one's current knowledge state and one's desired knowledge state. According to this proposal, curiosity is sparked by a *modest* information gap – that is, one large enough to indicate there's something to be learned, but not so large that satisfying one's curiosity seems unlikely. More recently, Dubey and Griffiths (2017, 2019) have proposed a rational theory of curiosity according to which curiosity functions to increase the value of an agent's current beliefs regarding the phenomena the agent expects to encounter in the future.

These forward-looking proposals have not been extended to explanation-seeking in particular, and thus have not considered functional consequences that may be unique to explanation search as compared to information search more generally. More importantly, the small handful of papers that have focused on explanation-seeking (as opposed to "fact-seeking") have investigated "backward-looking" triggers of information search, not an inference to anticipated (future) consequences. Both classic (e.g., Berlyne, 1950, 1966) and contemporary research (for a review, see Gottlieb, Oudeyer, Lopes, & Baranes, 2013) suggest that curiosity can be triggered by *novelty* and *surprise*, and these factors can indeed prompt explanation search, as well (Frazier, Gelman, & Wellman, 2009). Even the information gap theory, which has also been applied to explanation-seeking behavior (Mills, Sands, Rowles, & Campbell, 2019), has typically been tested by focusing on how confident participants are that they know the answer to a question (e.g., Gruber, Gelman, & Ranganath, 2014; Kang et al., 2009) or by how completely a question has been answered (Mills et al., 2019), both evaluations involving the relationship between the question and current beliefs, and not necessarily expectations about the consequences of seeking an answer.

Two theoretical proposals concerning explanation-seeking, both from philosophy, similarly focus on backward-looking triggers, though they emphasize the forward-looking goals these triggers serve. These accounts address "need for explanation"—the sense that a given event or phenomenon demands an explanation. Wong and Yudell (2015) propose that it is normatively correct to seek an explanation when a given event or phenomenon does not fit one's "map" of the world, a judgment which we refer to as *"map mismatch."* Relatedly, Grimm (2008) suggests that one experiences this sense of need for explanation when one can easily imagine an

[1] We thank an anonymous reviewer for highlighting this point.

[2] A body of research in social psychology has also considered the social and emotional consequences of engaging in explanation (Anderson, Krull, & Weiner, 1996; Malle, 2004). For example, an explanation could redirect blame, persuade, impress, or make one feel better about oneself. These are surely important psychological roles, but they are beyond the scope of the current paper, which focuses on the role of explanation in the process of inquiry, and hence on more epistemic functions.

Table 1
Potential triggers (with their primary functional consequences as applicable), and items used to measure each potential trigger in Studies 1–2 and Study 3.

Potential trigger	Item, Studies 1–2	Item, Study 3
Function: Learning about the explanandum		
Anticipated Learning	To what extent do you think the answer to this question would teach you something new?	To what extent do you think the answer to this question would teach you something new?
Learning Potential	Do you think there is something to be learned from the answer to this question (even if you yourself already know the answer)?	Do you think there is something to be learned from the answer to this question (even if you yourself already know the answer)?
Function: Learning for export		
Regularity	Do you think answering this question would help reveal a genuine pattern, structure, or regularity?	Do you think answering this question would help reveal a genuine pattern, structure, or regularity?
Future Utility	To what extent would knowing the answer to this question be useful to you in the future?	To what extent would knowing the answer to this question be useful to you in the future?
Simplicity (Reverse-Scored Complexity)	Do you think the answer to this question is likely to be simple or complex?	Do you think the answer to this question is likely to be simple or complex?
Breadth	Do you think the answer to this question is narrow (only applies to what is being explained) or broad (also applies to other similar cases)?	Do you think the answer to this question is narrow (only applies to what this question is asking about) or broad (also applies to other similar cases)?
Function: Learning from others' domain-based expertise		
Expertise	Do you think that answering this question correctly requires special expertise in some domain?	Do you think that answering this question correctly requires special expertise in some domain?
Complexity (Reverse-Scored Simplicity)	Do you think the answer to this question is likely to be simple or complex?	Do you think the answer to this question is likely to be simple or complex?
Backward-looking triggers from prior research		
Surprise	To what extent do you think it is surprising that *[premise]*?	Is this question asking about something surprising?
Novelty	How novel for you is the claim that *[premise]*?	Is this question asking about something that you find novel or new?
Information Gap*	How confident are you that you know the answer to this question?	How confident are you that you know the answer to this question?
Map Mismatch*	How well does the claim that *[premise]* fit in with your current beliefs about the world?	Does what this question is asking about fit in with your current beliefs about the world?
Fact-and-Foil	How easily can you imagine a world in which it is not the case that *[premise]*?	
Additional triggers		
Negative Valence*	To what extent do you think the main claim of the question (that *[premise])* is negative, positive, or neutral?	Is this question asking about something negative or positive?

* *Reverse-scored.*

alternative way the world could have been—that is, there is a salient contrast between *fact* (what actually occurred) and *foil* (what could have occurred instead). Both judgements involve an evaluation of the relationship between an event or phenomenon and prior beliefs (and are thus backward-looking), but the authors highlight their connection to subsequent epistemic consequences. For example, for Grimm, on receiving an explanation that differentiates between the fact and the foil, "the proto-understanding in terms of which we experience the world would, optimistically, be approximating ever more closely the various ways the world actually might be" (Grimm, 2008, p. 495).

In sum, prior research on question-asking has considered the expected consequences of pursuing information, but almost never in the context of explanation-seeking; research on fact-seeking and explanation-seeking curiosity, by contrast, has almost always relied on backward-looking cues (namely surprise, novelty, information gap, map mismatch, and fact-and-foil; see Table 1). In the following section, we bring a forward-looking, functional approach to the forefront, focusing in particular on the functional consequences of pursuing explanations.

1.2. Functional consequences of explanation search

In this section, we consider the functional consequences of seeking an explanation, and we derive candidate triggers of ESC on the basis of these consequences. Our guiding idea is that (a) ESC triggers explanation search, and (b) explanation search has a distinctive profile of downstream consequences. Under these assumptions, if ESC is well-calibrated with respect to the downstream consequences of explanation search, we would expect ESC to be triggered precisely when a learner judges it likely that these consequences will come about. In this way, identifying the functional consequences of engaging in explanation search can be used to generate testable hypotheses about potential triggers of ESC.

In the following sections, we review evidence to support three functional consequences of explanation search: learning about the explanandum (what is being explained), attaining information that is "exportable" beyond the explanandum, and learning from the domain-based expertise of others. We derive multiple candidate triggers from each function, as summarized in Table 1.

1.2.1. Learning about the explanandum

The first functional consequence of explanation search is obvious but important: explanation search, especially when successful, has the consequence of providing information about the explanandum (that is, the fact or event being explained), and in particular about how or why it is the case. For example, the explanation answering the question "Why is rust red?" will provide information about the properties of rust, and the inquirer is therefore likely to learn something new about the target of the explanation (rust, redness, and associated causal mechanisms). Learning about the explanandum can also occur when an individual "self-explains," even in the absence of feedback (Chi, De Leeuw, Chiu, & Lavancher, 1994). These findings motivate predictions about the triggers of ESC: ESC should be triggered whenever a person anticipates learning something new about the explanandum by engaging in explanation search (an idea that we refer to as *anticipated learning*, Table 1, Row 1), or when the explanation is expected to support learning, independent of the person's existing knowledge (an idea that we refer to as *learning potential,* Table 1, Row 2).

This prediction aligns with previous research on information search and question-asking. Oaksford and Chater (1994) formalized a measure of "expected information gain" that identifies the evidence one should collect to maximally reduce uncertainty in the space of candidate hypotheses. Since the introduction of this forward-looking measure, several studies have shown that adults' exploration is sensitive to expected information gain and thus the future value of information (e.g., Bramley, Lagnado, & Speekenbrink, 2015; Coenen, Rehder, & Gureckis, 2015; Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003). Furthermore, adults can evaluate (but not always generate) maximally informative questions in a Battleship-like game, with the goal of finding the location of several concealed ships on a game board (Rothe, Lake, & Gureckis, 2018). Even preschoolers can select the more effective of two questions for solving a particular problem as measured by expected information gain (Ruggeri, Sim, & Xu, 2017), and older children can often generate such questions themselves (Ruggeri & Lombrozo, 2015). These studies suggest that anticipated learning drives the evaluation of questions in tasks with well-defined goals; no prior work (to our knowledge) has investigated whether it affects the experience of curiosity about why or how something is the case.

1.2.2. Learning for export

Seeking explanations not only supports learning about the explanandum: it can also support learning that is "exportable" to other situations (Lombrozo & Carey, 2006). For example, an explanation for why some people are allergic to cats is likely to reveal general facts about the immune system and individual differences that are applicable well beyond the explained example. Indeed, self-explanation prompts are often beneficial (relative to other tasks) in helping students *transfer* the principle or procedure from explained problems to novel problems (Aleven & Koedinger, 2002; Atkinson, Renkl, & Merrill, 2003) or apply what they've learned to make novel inferences (Chi et al., 1994). Evidence suggests that these consequences are not incidental features of explanation-seeking, but rather arise because generalizability is among the criteria for what makes an explanation good. For instance, explanatory claims are judged more appropriate when they identify relationships that are "stable" in the sense that they hold despite variation in background circumstances, which should generally track exportability to new contexts (Vasilyeva, Blanchard, & Lombrozo, 2018; see also Blanchard, Vasilyeva, & Lombrozo, 2018).

These ideas motivate the prediction that ESC should be triggered by evidence that a to-be-received explanation will be exportable and applicable, as pursuing explanations under these conditions should increase the odds of obtaining good (i.e., exportable) explanations. More precisely, we might expect that people will be more curious about the answer to an explanation-seeking question when they expect it to pick out a genuine pattern or *regularity* that extends beyond the explanandum (Table 1, Row 3), and when they expect it to have high *future utility* (Table 1, Row 4). Dubey and Griffiths (2017, 2019) similarly predict effects of expected future utility on curiosity, but this prediction has not been evaluated in an explanation-seeking context.

Considerations regaring export also motivate more specific predictions concerning the roles of simplicity and breadth as triggers of ESC. Research suggests that all else being equal, people prefer explanations that are simple and broad (see Lombrozo, 2016 for a review). For example, when prompted to explain (rather than perform a control task), children and adults are more likely to discover patterns that are broad in scope (Walker, Lombrozo, Williams, Rafferty, & Gopnik, 2017; Williams & Lombrozo, 2010; Williams, Lombrozo, & Rehder, 2013) and simple in the sense that they appeal to few features (Kon & Lombrozo, in prep; Walker, Bonawitz, & Lombrozo, 2017). Importantly, such patterns may be more exportable than narrow and complex alternatives: an explanation that is broader in scope will necessarily explain *more* observations (Lombrozo & Carey, 2006), and simpler explanations may be more likely to generalize and support future communication and intervention (Pacer & Lombrozo, 2017). Correspondingly, we might expect ESC to be triggered when explanation-seeking is anticipated to help the learner identify a pattern that is expected to be *simple* (Table 1, Row 5) and *broad* (Table 1, Row 6), as these are the kinds of patterns that explanation-seeking seems to privilege.[3]

1.2.3. Learning from others' domain-based expertise

The preceding considerations (learning about the explanandum and learning for export) are likely to characterize explanation search quite broadly, whether the learner attempts to generate an explanation herself, engages in exploration to uncover an explanation, or asks a knowledgeable informant. However, this third method of explanation search—asking a knowledgeable informant—might introduce additional factors that could trigger ESC. A great deal of learning involves the testimony of others (Harris,

[3] It is unclear exactly how a learner would generate expectations about the simplicity or breadth of an explanation they have not yet received. However, there is already evidence that both children and adults have systematic expectations about the simplicity/complexity of causal mechanisms (Kominsky et al., 2017), as well as theoretical arguments to the effect that if pursuing an explanation is expected to have high epistemic utility (because it would be a good explanation if true), this could support its pursuit (Nyrup, 2015).

2012; Harris & Koenig, 2006), and testimony can often communicate content that is difficult to uncover through direct observation or exploration, such as causal structure or unobservable mechanisms. In fact, there's evidence that even young children will explore observable entities directly, but instead *ask questions* about unobservable entities (Fitneva, Lam, & Dunfield, 2013). Moreover, there is good evidence that people are adept at exploiting the "division of cognitive labor" (e.g., Keil, 2003; Keil, Stein, Webb, Billings, & Rozenblit, 2008; Lutz & Keil, 2002), deferring to appropriate experts in a given domain depending on what they want to know.

While deference to experts is likely to characterize question-asking in general, we might expect domain-based expertise to be especially relevant when it comes to explanation-seeking questions, given lay beliefs that domains are united by distinct patterns of causes and effects (Strickland, Silver, & Keil, 2017). In addition, there is evidence that people's willingness to defer to others to answer questions within a given domain is related to their attributions of *explanatory understanding*: attributions that another person *understands why* something in that domain occurred more closely track participants' willingness to defer to that person than do attributions that the person *knows why* (Wilkenfeld, Plunkett, & Lombrozo, 2016). Furthermore, adults' judgments of the quality of an explanation are predicted by their judgments of whether the explanation was written by an expert and their judgments of the explanation's complexity (Zemla, Sloman, Bechlivanidis, & Lagnado, 2017). Finally, children and adults are more likely to ask for help using or fixing a device that is judged to have a complex mechanism than a device that is judged to have a simple mechanism (Kominsky, Zamm, & Keil, 2017).

On our forward-looking approach, expectations about the necessity or value of consulting others could itself drive explanation-seeking curiosity. Specifically, if one way to pursue ESC is by asking knowledgeable others, then we might expect ESC to be triggered when *expertise* in some domain is needed to answer a given question (Table 1, Row 7), and/or when the answer is anticipated to be *complex* (Table 1, Row 8). Under these conditions, the learner might not expect to discover the answer on their own, but explanation-seeking directed towards domain-based experts can exploit the division of cognitive labor and support effective learning. While prior work shows that children and adults have the capacity to form expectations about expertise and complexity (e.g., Keil, Stein, Webb, Billings, & Rozenblit, 2008; Kominsky, Zamm, & Keil, 2017; Lutz & Keil, 2002), research has not investigated whether these expectations guide ESC.

1.3. The present research

Having articulated several potential triggers of ESC based on the downstream consequences of explanation seeking, and having reviewed several backward-looking triggers from prior work on question-asking, curiosity, and the need for explanation, we next turn to an empirical evaluation.

As described above, several of the potential triggers of ESC that we test have received previous empirical support, primarily for their role in driving information search more broadly (surprise, novelty, information gap, anticipated learning, and future utility). With the exception of surprise (Frazier et al., 2009) and information gap (Mills et al., 2019), however, prior work has not evaluated these factors as triggers of *explanation*-seeking curiosity or *explanation* search, in particular. Thus, one contribution of our work is to evaluate whether these candidate triggers indeed prompt ESC and explanation search. More importantly, however, our functional approach to ESC motivates a set of potential triggers that have not (to our knowledge) been tested as drivers of general curiosity and information search *or* of ESC and explanation search: expected simplicity/complexity, expected breadth, learning potential, and expected regularity. Additionally, while the literature on "need for explanation" has proposed normative standards for when an event *should* demand explanation, these standards (map mismatch, fact-and-foil) have not been empirically evaluated as triggers of ESC. A second contribution of our research is thus to evaluate whether these novel triggers prompt ESC and explanation search. Moreover, by testing all hypothesized triggers simultaneously, we can identify the relative contribution of each, and we can uncover patterns of shared variance that might indicate a meaningful, latent structure. These are important advantages over a more piece-meal approach focusing on the role of a single trigger.

Importantly, we can differentiate two hypotheses about the factors that drive ESC and explanation search. First, it is possible that a learner need not consider forward-looking triggers (such as anticipated learning and future utility) to successfully seek explanations in a way that achieves the downstream consequences of explanation. This is because backward-looking triggers, such as surprise and novelty, may be good (enough) proximal cues to the downstream consequences of explanation, obviating the need for a learner to *also* form expectations about anticipated learning and export. This hypothesis motivates the prediction that forward-looking triggers should play little to no role in predicting explanation-seeking curiosity once backward-looking triggers are considered. However, it is also possible that these forward-looking triggers have some added value for a learner above and beyond the backward-looking triggers, in which case both forward- and backward-looking triggers should predict a learner's experience of ESC. In the present research, we test these two hypotheses.

Methodologically, we also address a common limitation of previous research on explanation-seeking behavior. As spontaneous explanation-seeking is, by its very nature, difficult to capture in the lab, research has typically been motivated by a small number of realistic cases or relied on highly artificial contexts in the lab. The present research circumvents this challenge by drawing upon a novel source of data: large-scale on-line databases of user-generated questions that allow us to better capture explanation search in the course of everyday life. Specifically, we sample why-questions from on-line forums (e.g., "Why aren't clouds a uniform sheet of vapor in the sky?"), and we use these as stimuli in studies that prompt participants to report their curiosity about the explanandum (e.g., about why clouds aren't a uniform sheet of vapor in the sky) and to evaluate the question and explanandum along a variety of dimensions designed to track hypothesized triggers of ESC. This approach is correlational, which limits our ability to make claims about whether the hypothesized triggers *cause* a learner to experience ESC. However, the strengths of this approach—capturing a range of real-world explanation-seeking and testing many potential triggers of ESC in a single set of studies—offset this drawback.

In three studies, we address the following questions: (1) How well do our hypothesized triggers of ESC (Table 1) predict both explanation-seeking curiosity and explanation search itself? (2) Do forward-looking triggers contribute to the prediction of ESC above and beyond backward-looking triggers? (3) Is there some meaningful *structure* to the predictors of ESC (i.e., in terms of latent factors) that sheds light on function(s) of ESC? (4) How do the potential triggers of ESC compare to the potential triggers of *fact-seeking curiosity*?

In Study 1, we address the first and second questions by asking participants to rate ESC for a series of "why" questions, along with measures for the potential triggers of ESC identified in the prior sections. Additionally, we ask participants to select a subset of questions for which they would like to receive answers, thus allowing us to assess actual explanation-seeking behavior. In Study 2, we address the second and third questions by performing exploratory factor analysis on the potential trigger measures, using a stimulus set representing a broader sample of possible "why" questions. Finally, in Study 3, we answer the fourth question by comparing the triggers of ESC to the triggers of fact-seeking curiosity (FSC), using fact-seeking questions from previous research on curiosity (Dubey & Griffiths, 2017, 2019; Kang et al., 2009) in Study 3a, and using a novel set of both explanation-seeking and fact-seeking questions in Study 3b.

Broadly, these studies advance our understanding of ESC, both in terms of its relationship to other forms of curiosity and its potential triggers. Most notably, they highlight the importance of forward-looking considerations in guiding ESC above and beyond backward-looking considerations, and they shed light on the functional role of ESC and explanation seeking within the process of inquiry.

2. Study 1

In Study 1, we present participants with several questions and ask them to rate how curious they are about the answer to each question. Additionally, they rate each question on 13 potential trigger items, based on the potential triggers of ESC identified in the Introduction (see Table 1). We additionally include "negative valence" (Table 1, Row 13), a trigger motivated by work in attribution.[4] Subsequently, participants are asked to select the questions for which they would *most* like to receive an answer, thus providing the opportunity to seek actual explanations. This design allows us to assess whether ratings of ESC predict actual explanation search, validating our measure of ESC. In addition, we can test whether the 13 potential triggers identified in the Introduction do, in fact, predict ESC and explanation search. Finally, we investigate whether forward-looking triggers increase our ability to predict ESC beyond backward-looking triggers.

2.1. Method

2.1.1. Participants

Participants in all studies were recruited from Amazon Mechanical Turk and paid at the rate of $7.50 per hour, based on the expected average time to complete the task (20 min for Study 1 and 16 min for Studies 2–3). All participants provided informed consent under a protocol approved by the Institutional Review Board at University of California, Berkeley (Studies 1 and 2) or a protocol approved by the Institutional Review Board at Princeton University (Study 3). Participation in Studies 1 and 2 was restricted to users in the United States with a minimum approval rating of 95% based on at least 50 previous Mechanical Turk tasks.

The sample size for Study 1 was determined using the "simr" package for R (Green & MacLeod, 2016), simulating data based on a similar study from Liquin and Lombrozo (2018). This analysis suggested 80% power to detect an effect of ESC on explanation search with sample sizes between 150 and 200 participants. Our final sample for Study 1 included 195 adults (97 males and 98 females) ranging from 19 to 78 years of age ($M_{age} = 35$, $SD = 11$). Twenty-five additional participants were excluded for failing an attention check (explained below).

2.1.2. Materials

Forty "why" questions were sampled from the Reddit *Explain Like I'm Five* (ELI5) webpage (www.reddit.com/r/explainlikeimfive). On this page, Reddit users submit questions and other users provide easy-to-understand explanations. The questions were randomly sampled (from all questions posted between September and December 2017) and edited lightly for grammar and readability. Sample questions include: "Why are most pencils painted yellow? Why not a different color?"; "Why do ice cubes crackle when liquid is poured on them?"; and "Why does the bowl always get hotter than the food in the microwave?" For each question, we selected an answer provided by another user on the basis of its succinctness and accuracy; these were also edited lightly for grammar and readability. (See https://osf.io/mh8xf/ for a complete set of questions and answers.)

2.1.3. Procedure

Participants were randomly assigned to eight questions from the ELI5 stimulus set. For each question, participants were asked to rate their curiosity in response to that item: "Consider the following: *[that ice cubes crackle when liquid is poured on them]*. How curious are you about why this is the case?" This rating was completed on a seven-point scale, with the endpoints marked "Not at all curious

[4] Prior work on attribution shows that people tend to generate attributions for outcomes with negative emotional valence (Weiner, 1985; P. T. Wong & Weiner, 1981)—for example, a person is more likely to explain a failing grade on a test than a passing grade. While this potential trigger is not obviously forward- or backward-looking in nature, we include it due to its potential relevance to explanation search.

about why" and "Very curious about why." Additionally, participants rated each question on 13 other measures corresponding to the 13 potential triggers of ESC (with simplicity and complexity collapsed to a single dimension), summarized in Table 1. These 14 judgments (including ESC) were presented in a random order.

After completing these ratings, participants completed a short distractor task, which involved performing simple arithmetic problems. This task doubled as an attention check, as the task required participants to remember what they had seen previously (the answer to each arithmetic problem was an argument in the next arithmetic problem). Participants who made more than two errors were excluded.[5]

After the distractor task, participants were presented with the eight questions they had rated, and were instructed, "You get to see the answers to three of these questions. Please choose the three questions that you want answered." Participants read the explanations to the three questions they selected, after which they were asked to provide their age and gender (male, female, other, or prefer not to specify).

2.2. Results

2.2.1. Data preparation

Across all studies, listwise deletion was used to eliminate missing data for any of the potential trigger measures or ESC. Unless otherwise specified, all variables were z-scored to facilitate the interpretation of regression model parameters across studies. Descriptive statistics for curiosity and all potential trigger measures (across all studies) are reported in the supplementary material.

2.2.2. ESC and explanation search

First, we tested whether explanation search (i.e., revealing the answer to a question in the final phase of the study) could be predicted by ESC ratings. A generalized linear mixed model was fit to the answer reveal data (using the R package "lme4," Bates, Mächler, Bolker, & Walker, 2015, with a logit link function). The model contained ESC as a fixed effect, with random intercepts for participant and item.[6] This model was compared to a null model, which contained no fixed effects, using a likelihood ratio test. This analysis revealed a significant effect of ESC, OR = 1.69, 95% CI [1.51, 1.90], $\chi^2(1) = 84.94$, $p < 0.001$. This indicates that our measure of ESC predicts explanation search, and thus supports the validity of our measure of ESC.

2.2.3. Triggers of ESC

Next, we tested whether the 13 potential triggers predicted ESC. To test how well each trigger predicted ESC accounting for all other measures, we fit a regression model to the rating data. The potential triggers were fit as fixed effects, and random intercepts were fit for participant and item. The contribution of each potential trigger to the model was calculated by comparing the full model to a reduced model excluding that measure, using likelihood ratio tests. Seven of the 13 potential triggers made a significant contribution to the model predicting ESC, holding all other potential triggers fixed at their mean values: anticipated learning, learning potential, future utility, expertise, surprise, information gap,[7] and map mismatch (see Fig. 1; numeric values for all regression coefficients and 95% confidence intervals depicted in Fig. 1 are provided in the supplementary material).

We also investigated whether the 13 potential triggers were related to explanation search. Fitting a generalized linear mixed-effects model to the answer reveal data showed that anticipated learning, OR = 1.44, 95% CI [1.19, 1.71], $\chi^2(1) = 16.66$, $p < 0.001$, surprise, OR = 1.17, 95% CI [1.01, 1.37], $\chi^2(1) = 4.80$, $p = 0.03$, and information gap, OR = 1.26, 95% CI [1.10, 1.45], $\chi^2(1) = 11.58$, $p < 0.001$, made a significant contribution to the full model above and beyond the other potential trigger measures. The remaining measures did not reach statistical significance.

Because many of the 13 potential trigger measures were modestly correlated with each other, the failure of some measures to reach significance could be due to their shared variance, rather than their lack of relation to ESC. To address this, we also tested to what extent each measure predicted ESC and explanation search in isolation from all other measures (analyses in supplementary material). As expected, 10 of the 13 measures were significant predictors of ESC in isolation (all but fact-and-foil, map mismatch, and

[5] While the primary function of this attention check was to ensure that participants were reading instructions and paying attention, it could be argued that working memory and arithmetic abilities are not diagnostic of attentive completion of this study. As a result, we also conducted all analyses without any exclusions. The pattern of results reported remained unchanged.

[6] Following the recommendation of Barr, Levy, Scheepers, and Tily (2013) and an anonymous reviewer, we attempted to fit all models with by-participant and by-item random slopes, in addition to random intercepts. In nearly all cases, the model with random slopes and random intercepts resulted in singular fit, indicating that the random effects structure was too complex to be supported by the data. Following the recommendation of Singmann and Kellen (2019), we subsequently simplified the models, first eliminating by-item random slopes, then all random slopes. In most cases, only the fully simplified (random-intercept only) models converged, so we report these models. In cases where more complex models converged, the parameter estimates for fixed effects were comparable to the random-intercept only models. All alternative models are included in our analysis scripts at https://osf.io/mh8xf/.

[7] We also tested whether the quadratic effect of information gap on curiosity was significant, given the proposed "U-shaped curve" relating the magnitude of the information gap to curiosity (Loewenstein, 1994). To do so, we fit a multi-level model with random intercepts for participant and item, and fixed linear and quadratic terms for information gap. This quadratic model was compared to a model excluding the quadratic term using a likelihood ratio test. The results of this test revealed a significant quadratic effect, $\chi^2(1) = 6.76$, $p = 0.009$. However, inspection of the model coefficients reveals that the magnitude of this effect was quite small, $\beta = -0.07$, 95% CI $[-0.12, -0.02]$. Therefore, we take the linear effect alone to be a reasonable approximation of the effect of an information gap on curiosity in all subsequent analyses.

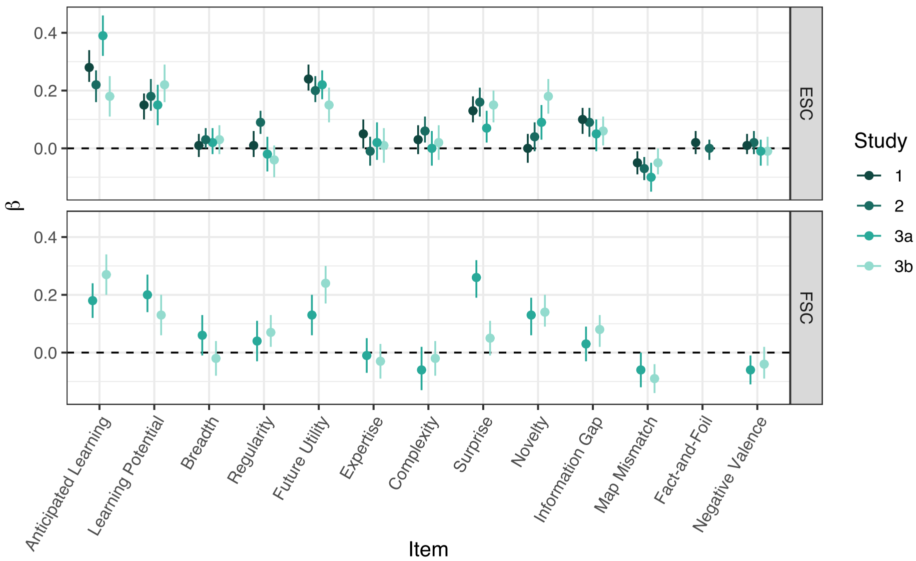

Fig. 1. Standardized regression coefficients (with 95% CIs) predicting ESC/FSC for each potential trigger.

negative valence), and nine of the 13 potential triggers were significant predictors of explanation search (all but breadth, fact-and-foil, map mismatch, and negative valence).

2.2.4. Forward-looking and backward-looking triggers

Finally, we tested whether forward-looking triggers of ESC significantly contributed to the prediction of ESC and explanation search, over and above the predictive power afforded by backward-looking triggers. In other words, do forward-looking expectations about the functional consequences of engaging in explanation search capture variance in ESC or explanation search that was not already captured by backward-looking considerations (such as surprise and information gap)? We fit two multi-level models to the ESC rating data. The first model contained all backward-looking triggers as fixed effects: surprise, novelty, information gap, map mismatch, and fact-and-foil. The second model also contained these, plus fixed effects for simplicity/complexity, breadth, anticipated learning, learning potential, regularity, future utility, and expertise (negative valence was excluded from this analysis). Both models included random intercepts for participant and for item. The two models were compared using a likelihood ratio test. This analysis suggested strongly that the forward-looking triggers of explanation search significantly contributed to the prediction of ESC above and beyond backward-looking triggers, $\chi^2(7) = 578.19, p < 0.001$. This analysis was repeated for the answer reveal data (removing random intercepts for participant, as all participants selected exactly three answers to reveal), and similarly, the forward-looking triggers significantly increased the model fit, $\chi^2(7) = 41.89, p < 0.001$. According to Nakagawa and Schielzeth's (2013) measure of R^2 for generalized linear mixed effects models, adding the forward-looking potential triggers increased the marginal R^2 (measuring only the percentage of variance explained by the fixed effects of the models) from 21% to 51% for the prediction of ESC and from 5% to 9% for the prediction of answer reveal. Thus, while backward-looking triggers may serve as proximal cues to the downstream consequences of explanation seeking, our forward-looking triggers play a role in predicting ESC and explanation-seeking that is not well captured by these backward-looking triggers.

2.3. Discussion

These results suggest that our measure of ESC predicts actual explanation search, and that many of the potential triggers of ESC identified in the Introduction predict both ESC and explanation search. Most strikingly, the strongest three predictors of ESC—anticipated learning, learning potential, and future utility—align with forward-looking goals of learning about the explanandum and learning for export. Other predictors of ESC—information gap, surprise, and map mismatch—are backward-looking, and are more familiar from previous research on curiosity and need for explanation. In Studies 2–3 we consider the relationships among triggers and their associations with ESC more systematically, but for now we note that the best predictors of ESC are "forward-looking" as opposed to "backward-looking" in that they concern the *potential benefits* of explaining the explanandum, rather than the relationship between the explanandum and *current beliefs*. Furthermore, forward-looking triggers predict ESC and explanation search above and beyond backward-looking triggers, suggesting that there may be value to considering expectations about the future when seeking explanations that goes beyond attention to novelty, surprise, and other backward-looking considerations.

While some potential triggers failed to significantly predict ESC or explanation search, even in isolation (see supplementary material), they may not be irrelevant to ESC in all contexts. In particular, the weak association between negative valence and ESC is plausibly a consequence of the nature of our stimuli: while research in attribution theory tends to focus on personally-relevant events, the questions used in the present research were not particularly personally relevant, nor did they seek to explain success or failure at achieving a given goal. We also failed to find a positive association between simplicity and ESC; instead, ESC was higher for questions with answers that were expected to be more *complex*. Even if some forms of explanatory simplicity are valued (Blanchard, Lombrozo, & Nichols, 2018; Bonawitz & Lombrozo, 2012; Lombrozo, 2007; Pacer & Lombrozo, 2017; Read & Marcus-Newhall, 1993), this could be because complex explanations can make the explanandum appear more "necessary" (Zemla et al., 2017) or probable (Johnson, Valenti, & Keil, 2019), or for the reasons that led us to predict an independent effect of complexity: a complex explanation might require greater reliance on experts, and reveal more about the explanandum.

Somewhat surprisingly, the fact-and-foil and map mismatch accounts (Grimm, 2008; W. Wong & Yudell, 2015), the two most fully developed accounts of "need for explanation" to date, fared quite poorly in predicting ESC. Fact-and-foil was not a significant predictor of ESC, and map mismatch was a *negative* predictor of ESC in the simultaneous regression model, indicating that a higher mismatch between one's map of the world and a current observation predicted lower levels of curiosity. Again, however, we do not take our results as evidence that these considerations are irrelevant to ESC. Here, several caveats are in order: First, these theories of explanation search were not necessarily intended to be descriptive psychological models, and thus people may in fact *not* use these considerations as cues to ESC, even if (perhaps) they should. Second, these accounts are plausibly directed at differentiating explanation-seeking questions that arise from those that do not even arise, while in this study, we focused on variation in ESC among questions that had in fact been asked. Put differently, our stimuli were restricted to "why" questions that already met some threshold for "need for explanation," and it could be that the fact-and-foil and map mismatch accounts identify factors that differentiate what falls under versus over this threshold, rather than accounting for variation in strength of ESC well above this threshold. In Study 2, we address this concern.

3. Study 2

In Study 2, we replicate Study 1 with a broader set of "why" questions drawn from two sources. The *Explain Like I'm Five* question set likely reflects a very limited range of all possible questions, as these questions were posted on a website in search of answers. To truly discover the sources of ESC, we would ideally probe such "asked" questions alongside "unasked" questions—that is, questions that haven't been posted on a webpage specifically designed to answer questions, and that reflect questions that people *would not* normally be compelled to ask. This is precisely what we do in Study 2 by constructing a set of "unasked" questions in addition to a new set of *Explain Like I'm Five* questions. This allows us to test the 13 potential triggers more fairly, using a broader set of potential "why" questions. The addition of this question set also increases the probability that any detected triggers of ESC distinguish ESC from its absence (rather than distinguishing high levels of ESC from moderate levels of ESC), and it addresses the potential worry that questions *posed to others* (on an on-line forum) could inflate the importance of expertise as a predictor of ESC.

Additionally, as suggested by the modest correlations between several of the 13 potential triggers of explanation search, it is likely that these judgments reflect some underlying structure. If we assume the potential trigger ratings are responded to on the basis of several latent variables, we can use factor analysis to extract these latent variables. These latent variables, in turn, might provide new insights about both the triggers of ESC and how they map onto the functional consequences of explanation search. In Study 2, we explore this potential structure, finding three latent variables that predict ESC.

3.1. Method

3.1.1. Participants

Participants were 187 adults (80 males, 106 females, and 1 other) ranging from 19 to 70 years of age (M_{age} = 36, SD = 12). Fourteen additional participants were excluded for failing an attention check (explained below). The sample size was planned to match the sample size of Study 1.

3.1.2. Materials

Forty new "why" questions (not used in Study 1) were sampled from the Reddit *Explain Like I'm Five* webpage (www.reddit.com/r/explainlikeimfive). Questions were edited lightly for grammar and readability. (See https://osf.io/mh8xf/ for the complete set of the stimuli.)

Since ELI5 is comprised only of questions that users actually asked (and thus these questions likely already surpass some threshold of curiosity beyond which a question is posed), we also sought "unasked questions" from ConceptNet, an open source semantic network (http://conceptnet.io; Speer, Chin, & Havasi, 2017). This semantic network contains words and phrases, which are linked to other words and phrases by the relations they hold to each other. For example, the words "lizard" and "garden" are linked by the relation "location of." Subsequently, a "why" question can be constructed from these two words and their relation: "Why are lizards found in gardens?" To constrain our construction of questions, we focused solely on the domain of non-human animals. We selected ten animals (lizard, dog, beaver, camel, horse, bear, dolphin, duck, deer, and hamster) and four relation types (is a type of, is capable of, location of, and types of), then constructed 40 why questions based on the entries under these animals and relations in ConceptNet. Sample questions include: "Why are dogs a type of canine?"; "Why are dromedaries a type of camel?"; and "Why are bears capable of fishing for salmon?"

3.1.3. Procedure

Participants were assigned to eight questions from either the ELI5 ($N = 96$) or ConceptNet ($N = 91$) question set.[8] For the ConceptNet question set, participants were randomly assigned two questions from each relation type.

First, as in Study 1, participants rated each item on ESC and the 13 other potential trigger measures. After completing these ratings for each of the eight questions, participants completed an attention check, in which they were asked to identify four of the questions they had read previously from four distractor questions. The distractor questions were four additional questions from ConceptNet or from ELI5 that were not included in the full question set. The attention check was scored for accuracy out of eight (one point for each question correctly selected and one point for each distractor question not selected), and participants who scored lower than six were excluded.

3.2. Results

3.2.1. Triggers of ESC

First, we tested the association between each potential trigger and ESC, accounting for the other potential triggers by holding them fixed at their mean values. We thus repeated the analysis used in Study 1, finding that 8 of the 13 potential triggers made a significant contribution to the model predicting ESC: anticipated learning, learning potential, regularity, future utility, complexity, surprise, information gap,[9] and map mismatch (see Fig. 1). These results largely replicate Study 1; as in Study 1, many forward-looking potential triggers (anticipated learning, learning potential, regularity, future utility, and complexity) were significant predictors of ESC even holding backward-looking triggers fixed. We also replicate the results of the individual regression analyses in the supplementary material.

3.2.2. Stimulus set

We were also interested in whether the stimulus set (ELI5 or ConceptNet) moderated the relationship between any potential trigger and ESC. We expanded upon the model reported above by adding regression coefficients for stimulus set and for the interactions between stimulus set and each potential trigger measure. The significance of each interaction was tested by comparing (using likelihood ratio tests) the full model to a reduced model excluding a single interaction term. The interactions between stimulus set and complexity, $\beta = -0.13$, 95% CI [-0.22, -0.03], $\chi^2(1) = 7.06$, $p = 0.008$, and stimulus set and information gap, $\beta = 0.10$, 95% CI [0.004, 0.19], $\chi^2(1) = 4.18$, $p = 0.04$, made a significant contribution to the full model. Complexity was a stronger predictor of ESC in the ConceptNet question set than the ELI5 question set, while information gap was a stronger predictor of ESC in the ELI5 question set than the ConceptNet question set. In general, this analysis suggests only minor differences between "asked" (ELI5) and "unasked" (ConceptNet) questions in the extent to which the potential triggers predict ESC.

3.2.3. Factor structure of predictors

Having replicated the basic pattern of results from Study 1, we next performed a factor analysis on the 13 potential triggers of ESC. Raw ratings on the 13 items were group mean centered on the basis of individual questions (e.g., the mean "learning potential" rating for the question "Why are dogs a type of canine?" was subtracted from all "learning potential" ratings in response to that question). This technique reduces the ICC, which in turn reduces the influence of clustering on parameter estimates (see Huang & Cornell, 2016). Initial eigenvalues for the first three factors were over one, and explained 29%, 14%, and 10% of the variance in ratings, respectively. The next three factors explained 7%, 7%, and 6% of the variance and had eigenvalues less than one. In accordance with the results of parallel analysis, we retained a three-factor solution.

An initial three-factor solution, using varimax rotation, was inspected. Two potential triggers—fact-and-foil and negative valence—were eliminated because they failed to load onto any factor with a factor loading above 0.4 or any several factors with loadings above 0.3. The final three-factor solution with varimax rotation, excluding these two items, explained 47% of the total variance in ratings. The three factors corresponded roughly to learning-related considerations (anticipated learning, learning potential, information gap, expertise, and complexity), export-related considerations (learning potential, regularity, future utility, and breadth), and backward-looking considerations (surprise and novelty). Though map mismatch loaded onto the backward-looking factor in the initial factor analysis, its factor loading did not exceed 0.4 in the final solution (once fact-and-foil and negative valence were removed). The factor loading matrix for this analysis is presented in Fig. 2.

3.2.4. Factor scores and ESC

Next, we extracted factor scores from the final three-factor solution (using Thomson's method; Thomson, 1951), for use in predicting ESC. The z-scored factor scores were simultaneously entered as fixed effects in a model predicting ESC, with random

[8] Participants for each stimulus set were in fact recruited separately, but the studies are reported here as one study, as methods and analyses were identical other than the stimulus set used.

[9] Again, we tested the quadratic effect of information gap on curiosity, using the same method as in Study 1. This revealed a significant quadratic effect, $\chi^2(1) = 13.65$, $p < 0.001$. However, the magnitude of this effect was again quite small, $\beta = -0.10$, 95% CI [-0.16, -0.05]. We thus take the linear effect alone to be a reasonable approximation of the effect of an information gap on curiosity in all subsequent analyses.

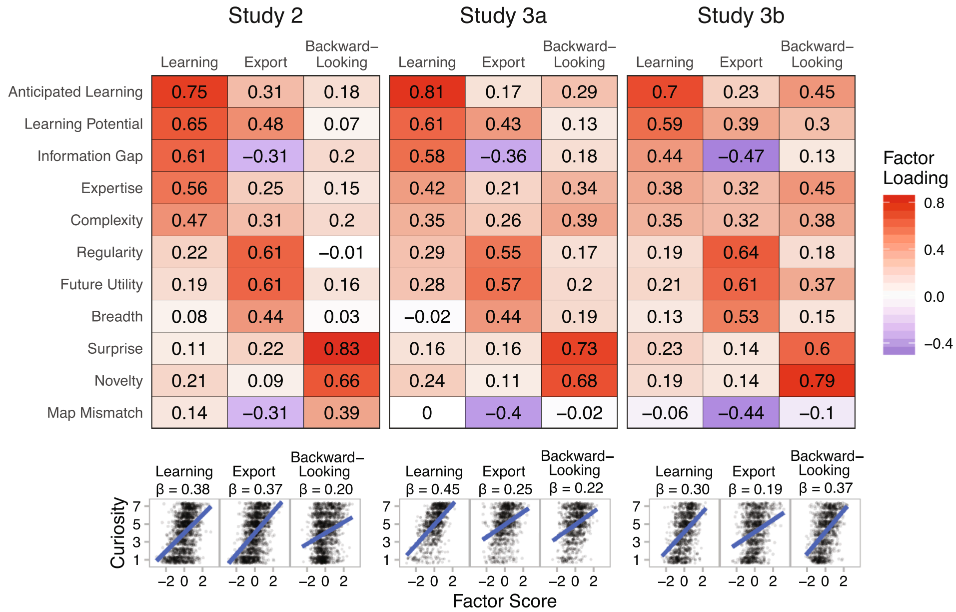

Fig. 2. Results of exploratory factor analysis on ratings in response to explanation-seeking questions. For Studies 3a and 3b factors are rearranged to facilitate comparison with Study 2 (Study 3a original order: learning, backward-looking, export; Study 3b original order: export, backward-looking, learning). Below, jittered scatterplots (with a point for each item rated by each participant) showing the relationship between extracted factor scores for each factor/study and ESC ratings, and the standardized regression coefficient relating these variables controlling for the nested structure of the data.

intercepts for participant and item.[10] The contribution of each factor to the prediction of ESC was calculated by comparing this model using likelihood ratio tests to three reduced models, each excluding one of the three factors. This analysis revealed that the factor scores for all three factors made a significant contribution to the prediction of ESC (see Fig. 2), learning: $\beta = 0.38$, 95% CI [0.35, 0.42], $\chi^2(1) = 386.31, p < 0.001$; export: $\beta = 0.37$, 95% CI [0.33, 0.41], $\chi^2(1) = 276.43, p < 0.001$; backward-looking: $\beta = 0.20$, 95% CI [0.16, 0.23], $\chi^2(1) = 124.95, p < 0.001$.

For each factor, we also fit a separate model to predict ESC ratings, with the factor score as a fixed effect and random intercepts for participant and item. Nakagawa and Schielzeth's (2013) measure of R^2 for generalized linear mixed effects models was used to compare the variance in ESC explained by each model. The marginal R^2 of the learning model was 0.20, the marginal R^2 of the export model was 0.18, and the marginal R^2 of the backward-looking model was 0.06. These findings suggest that the learning model explained the most variance in ESC, followed by the export model then the backward-looking model.

3.3. Discussion

Study 2 offers several insights about explanation-seeking curiosity and the mapping between the factors that trigger it and the functional consequences of explanation search. First, we replicated the results of Study 1 on a broader set of "why" questions that included both "asked" (ELI5) and "unasked" (ConceptNet) questions. Additionally, ratings of the 13 potential triggers of ESC corresponded to a three-factor structure, suggesting that the triggers of explanation search correspond to three latent variables, tracking forward-looking potential for learning, forward-looking potential for export, and backward-looking considerations. Notably, these first two factors correspond fairly well with the functional consequences of explanation search reviewed in the Introduction (with the triggers supporting "learning about the explanandum" and "learning from others' domain-based expertise" collapsed into a single "learning" factor). Information gap, which we initially classified as a backward-looking trigger, in fact patterned with forward-looking triggers concerning expectations about learning. While this item ("How confident are you that you know the answer to this question?") does concern prior knowledge rather than expectations about the future, and is thus backward-looking in nature, it may be the case that this backward-looking measure is in fact a highly reliable cue to anticipated learning. For example, some theories of curiosity suggest that an information gap leads to a feeling of deprivation, which results in curiosity and motivates subsequent

[10] Lastovicka and Thamodaran (1991), among others, have suggested that factor scores may not be appropriate for use in regression analysis because these scores are estimates rather than measurements, resulting in biased regression coefficient estimates. We report this method in the main text because it is familiar and easy to interpret, but we also present the results from an alternative method, using Structural Equation Modeling (SEM), in the supplementary material. These additional analyses are provided for all analyses for which factor scores are used across all studies, and produce largely consistent results with the factor score regression analysis.

learning (e.g., Jirout & Klahr, 2012). Interestingly, information gap may in fact be a *better* cue to expectations about learning than are surprise and novelty, as these measures load onto a separate factor and thus are less closely associated with anticipated learning.

Our backward-looking factor, which included measures from classic theories of curiosity (e.g., Berlyne, 1950, 1966; Berlyne & Frommer, 1966), captured the smallest portion of variance in people's judgments of ESC. Relatedly, the map mismatch account (W. Wong & Yudell, 2015) and the fact-and-foil account (Grimm, 2008), the principal theoretical accounts of need for explanation to date, were again fairly poor predictors of actual ESC judgments (replicating Study 1). Instead, ESC was best captured by triggers corresponding to learning and export: the extent to which an explanation might teach one something new, and the extent to which that information might be generalizable to other contexts in the future. This is surprising because it has been assumed in previous research on curiosity (whether implicitly or explicitly) that backward-looking considerations (e.g., surprise) are important drivers of curiosity and exploratory behavior *because* they are cues to important functional consequences, such as learning. As in Study 1, we show that forward-looking considerations (in the form of *expectations* about downstream functional consequences) may in fact be important correlates of curiosity above and beyond these backward-looking considerations.

Nonetheless, these results leave several questions unanswered. First, how well does the three-factor structure we found in this study extend to different sets of explanation-seeking questions? We test this in Studies 3a and 3b, conducting a close replication with 40 new questions from *Explain Like I'm Five* (Study 3a) and a far replication with 40 questions drawn from a new source, notably one for which the explananda are themselves novel for participants (Study 3b). Second, how do the potential triggers of *explanation-seeking curiosity* compare to the potential triggers of *fact-seeking curiosity*?

4. Study 3

In Study 3, we attempt to replicate the three-factor structure describing the potential triggers of ESC in Study 2. Additionally, we investigate whether the tested triggers of ESC are differentially relevant for the prediction of *fact-seeking curiosity* (FSC). We do so using two sets of questions as experimental stimuli. In Study 3a, we use a new set of ELI5 questions, and contrast these questions with fact-seeking trivia questions used in prior work (Dubey & Griffiths, 2017, 2019; Kang et al., 2009). In Study 3b, we use explanation-seeking and fact-seeking questions posted as comments in response to captioned images in an on-line forum. This latter set of stimuli has an additional notable feature: while the ELI5 questions and trivia questions largely ask about familiar, everyday objects or events (e.g., "Why do ice cubes crackle when liquid is poured on them?"), and the ConceptNet items probed familiar facts from common knowledge (e.g., "Why are dogs a type of canine?"), the questions used in Study 3b are in response to novel image stimuli. Thus, these questions provide a fairer test of the role of "backward-looking" considerations (e.g., map mismatch, surprise, and novelty) in predicting ESC, as novel stimuli have a much greater potential to invoke surprise, novelty, or a map mismatch than familiar stimuli.

Why might we expect differences in the predictors of ESC and FSC? Lowenstein (1994) drew upon research on explanation search to build his information gap theory of curiosity, and Schwitzgebel (1999) suggests that explanation-seeking curiosity could be either a different *type* of curiosity, or simply the same drive state directed towards a different type of information. However, some lines of research point towards differences in explanation-seeking and fact-seeking processes. First, the "explanation for export" account emphasizes the role of explanation in providing information that can be generalized to future contexts (Lombrozo & Carey, 2006). While this account does not address the contrast between explanations and facts, it might suggest that export should be *more* characteristic of explanations than of other types of information (including facts), and thus we might expect export-relevant potential triggers—such as future utility, breadth, and regularity—to be stronger predictors of ESC than of FSC.

Additionally, there is evidence for a special link between explanation and *deference to experts*, which may not exist to the same extent for simple facts. Wilkenfeld et al. (2016) presented participants with a series of hypothetical agents who possessed different depths of knowledge about a given causal connection, and asked the participants to evaluate the extent to which each agent *knew why* or *understood why* a relevant effect occurred. Relatively shallow levels of knowledge were sufficient for attributions of knowledge-why, but greater depths of knowledge were required for attributions of understanding-why. Notably, judgments of understanding-why (vs. knowledge-why) were also more strongly associated with a measure of deference to expertise: whether the agent would be a good person to consult for a related question in the same domain. This suggests a special link between understanding why something is the case (i.e., having an explanation) and expertise, which does not extend to mere knowledge. Therefore, in the context of this study, we might expect expertise-relevant potential triggers to be stronger predictors of ESC than of FSC.

4.1. Method

4.1.1. Participants

Participants in Study 3a were 234 adults (113 males, 119 females, and 2 other) ranging from 20 to 72 years of age (M_{age} = 36, SD = 10). Participants in Study 3b were 251 adults (109 males, 141 females, and 1 other) ranging from 19 to 73 years of age (M_{age} = 36, SD = 10). Across the two studies, fifty additional participants were excluded for failing attention checks (explained below), and six additional participants were excluded because their responses came from duplicate IP addresses.

The sample sizes of these studies were planned to match Study 2 after exclusions. An initial version of Study 3b contained a significant proportion of low-quality data, likely due to automated responding or "bots" on Amazon Mechanical Turk (Dennis, Goodson, & Pearson, 2018). To combat this, we added several attention checks and restricted participation to those in the United States with a minimum approval rating of 99% on at least 1000 previous Mechanical Turk tasks. We also included a Captcha

Verification in these studies. We preregistered the studies and analyses,[11] but some of the analyses reported below go beyond our preregistration. All preregistered analyses are presented in the supplementary material, and all departures and additions are explained and justified.

4.1.2. Materials

For Study 3a, forty new "why" questions (not used in Studies 1 or 2) were sampled from the Reddit *Explain Like I'm Five* webpage (www.reddit.com/r/explainlikeimfive). Questions were edited lightly for grammar and readability. The 40 trivia questions used by Kang et al. (2009) and Dubey and Griffiths (2017, 2019) were used as fact-seeking questions (FSQs).

For Study 3b, we used Google's BigQuery (bigquery.cloud.google.com) to extract comments on the r/pics subreddit (www.reddit.com/r/pics/) posted between January 2017 and June 2018 that began with the words "who," "what," "when," where," "why," or "how." All comments were posted directly in response to the original posts (which were images with captions), rather than in response to another user's comment. From this set of comments, we extracted posts that had at least one why-question comment and at least one who/what/when/where/how-question comment, then hand selected 40 posts with an explanation-seeking why-question and a fact-seeking who/what/when/where/how-question,[12] avoiding rhetorical questions, jokes, and inappropriate content. This produced a set of 40 explanation-seeking questions and 40 fact-seeking questions, posted in response to the same 40 images and captions. As an example, one image was a picture of two dogs on a snowy hill at sunset, with the caption "My mom took this at our cabin yesterday." The explanation-seeking question in response to this post asked, "Why is only one of your dogs on a leash and not the other?" and the fact-seeking question asked, "Whereabouts was this?" (See https://osf.io/mh8xf/ for a complete set of stimuli).

4.1.3. Procedure

Participants were randomly assigned to eight questions from either the ESQ set ($N_{3a} = 114$; $N_{3b} = 128$) or the FSQ set ($N_{3a} = 120$; $N_{3b} = 123$). For each question, participants responded to the prompt "How curious are you about the answer to this question?" on a seven-point scale, with the endpoints marked "Not at all curious" and "Very curious." Participants also rated each question on 12 potential trigger measures (fact-and-foil was excluded due to our inability to produce a version of this item that applied to FSQs), each on a seven-point scale. Several of these items were reworded from the items in Studies 1 and 2 so that they would apply to both explanation-seeking and fact-seeking questions (see Table 1). These 12 items and the curiosity measure were presented in a random order.

After completing these ratings for each of the eight questions, participants completed three attention checks. First, they were asked to type a given word contained in a JPEG image into a text box. Next, they were asked to identify four of the questions (Study 3a) or images (Study 3b) they had seen previously from four distractor questions/images. The distractor questions were four additional questions from ELI5 (for ESQs) or four additional fact-seeking trivia questions (for FSQs), and the distractor images were four additional images from other Reddit posts. This attention check was scored as in Study 2. Finally, participants solved an addition problem, for which the arguments were two random numbers between 1 and 20. Participants who did not successfully complete all three attention checks were excluded.

4.2. Results

4.2.1. Triggers of ESC and FSC

First, we tested whether the 12 potential triggers were differentially predictive of ESC and FSC. To do so, we fit a mixed-effects regression model predicting curiosity, with fixed effects for each potential trigger, question type (explanation-seeking vs. fact-seeking), and the interaction between each trigger and question type; and with random intercepts for participant and item. This full model was compared to a set of reduced models that each excluded one interaction term, using likelihood ratio tests. In Study 3a, anticipated learning was a stronger predictor of ESC than of FSC, $\beta = 0.21$, 95% CI [0.11, 0.31], $\chi^2(1) = 17.13$, $p < 0.001$, while surprise was a stronger predictor of FSC than of ESC, $\beta = -0.18$, 95% CI [−0.27, −0.10], $\chi^2(1) = 17.82$, $p < 0.001$. In Study 3b, surprise, $\beta = 0.09$, 95% CI [0.01, 0.17], $\chi^2(1) = 5.33$, $p = 0.02$, and learning potential, $\beta = 0.10$, 95% CI [0.002, 0.19], $\chi^2(1) = 3.97$, $p = 0.046$, were stronger predictors of ESC than of FSC. Regularity, $\beta = -0.12$, 95% CI [−0.20, −0.04], $\chi^2(1) = 8.68$, $p = 0.003$, and future utility, $\beta = -0.09$, 95% CI [−0.17, −0.002], $\chi^2(1) = 4.04$, $p = 0.04$, were stronger predictors of FSC than of ESC. Notably, these results are inconsistent across studies and do not align with our predictions; although ESC and FSC may diverge, the measures that predict one over the other may vary across settings.

Next, we fit a model predicting curiosity across conditions (collapsing across ESC and FSC), using the same analysis as in Studies 1 and 2. In Study 3a, eight of the 12 potential trigger measures made a unique contribution to the prediction of curiosity, holding all other measures fixed at their mean values: anticipated learning, $\beta = 0.26$, 95% CI [0.21, 0.30], learning potential, $\beta = 0.19$, 95% CI [0.15, 0.24], future utility, $\beta = 0.19$, 95% CI [0.14, 0.23], surprise, $\beta = 0.17$, 95% CI [0.13, 0.22], novelty, $\beta = 0.10$, 95% CI [0.06,

[11] The preregistration for Study 3a can be found at https://aspredicted.org/bx6bh.pdf and the preregistration for Study 3b can be found at https://aspredicted.org/8je27.pdf.

[12] "How" questions are often explanation-seeking in nature (e.g., "How does a bicycle work?"). For the purpose of this study, we exclusively selected *fact-seeking* "how" questions (e.g., "How tall is the Eiffel Tower?").

0.15], information gap, $\beta = 0.05$, 95% CI [0.01, 0.09],[13] map mismatch, $\beta = -0.08$, 95% CI [−0.12, −0.04], and negative valence, $\beta = -0.04$, 95% CI [−0.08, −0.01]. In Study 3b, all the same measures made unique significant contributions to the prediction of curiosity, with the exception of negative valence: anticipated learning, $\beta = 0.23$, 95% CI [0.18, 0.28], learning potential, $\beta = 0.18$, 95% CI [0.13, 0.23], future utility, $\beta = 0.19$, 95% CI [0.15, 0.23] surprise, $\beta = 0.10$, 95% CI [0.06, 0.14], novelty, $\beta = 0.16$, 95% CI [0.12, 0.20], information gap, $\beta = 0.08$, 95% CI [0.04, 0.11], and map mismatch, $\beta = -0.07$, 95% CI [−0.10, −0.04]. We also fit separate models for ESC and FSC to facilitate comparison with the previous studies, which are reported in Fig. 1 and the supplementary material. These results are largely consistent with the results of Studies 1 and 2. In particular, we again see that the strongest predictors of curiosity are forward-looking (future utility, anticipated learning, and learning potential), while the backward-looking measures emphasized in prior research on curiosity (information gap, surprise, novelty) play a smaller role.

4.2.2. Factor structure and ESC

We next sought to verify the factor structure found in Study 2 in response to explanation-seeking questions. We did so using two approaches. First, we used confirmatory factor analysis to test whether the factor solution from Study 2 provided a good fit to the data from Study 3a and Study 3b. This analysis is presented in the supplementary material. The confirmatory factor analyses did not meet standard thresholds of good fit, so we then turned to exploratory factor analysis to investigate the factor structure of the potential trigger measures in these two new datasets (as preregistered, and as recommended by Flora & Flake, 2017).

As in Study 2, raw ratings on the measures were group-mean centered within questions, to eliminate the influence of clustered data on parameter estimates. For Study 3a, initial eigenvalues for the first three factors were over one, and explained 33%, 13%, and 9% of the variance in ratings, respectively. Consistent with parallel analysis, we retained a three-factor solution. For Study 3b, initial eigenvalues for the first two factors were over one, and explained 37% and 13% of the variance in ratings, respectively. The third factor had eigenvalue 0.97 and explained 8% of the variance. Parallel analysis recommended a two-factor solution. Because three-factor solutions were used for all prior analyses, we inspected both two- and three-factor solutions; the three-factor solution produced better fit to the data than the two-factor solution; $TLI_{3\text{-factor}}$: 0.94, $RMSEA_{3\text{-factor}}$: 0.06, 95% CI [0.05, 0.07]; $TLI_{2\text{-factor}}$: 0.90, $RMSEA_{2\text{-factor}}$: 0.08, 95% CI [0.07, 0.09]. Additionally, the three-factor solution was more interpretable and would facilitate comparison to Studies 2 and 3a. Thus, we report the final two-factor solution in the supplementary material, but we retain the three-factor solution in the text that follows.

An initial three-factor solution, using varimax rotation, was inspected for each study. In both datasets, negative valence was eliminated because it failed to load onto any factor with a factor loading above 0.4 or on several factors with loadings above 0.3. The final three-factor solutions with varimax rotation, excluding this item, explained 44% of the total variance in ratings for Study 3a and 48% of the total variance in ratings for Study 3b. Notably, the structure of the latent variables was quite similar to that in Study 2 (see Fig. 2).

To quantitatively compare the results of these factor analyses to the Study 2 factor analysis, we estimated the correlation between the factor loadings for each potential trigger measure onto each factor in Study 2 and the corresponding factor loadings in Study 3a and Study 3b. The factor solutions were highly correlated between Study 2 and Study 3a, $r = 0.90$, $t(31) = 11.79$, $p < 0.001$, between Study 2 and Study 3b, $r = 0.82$, $t(31) = 8.03$, $p < 0.001$, and between Study 3a and Study 3b, $r = 0.94$, $t(31) = 16.03$, $p < 0.001$. These results suggest that there is significant overlap between the factor solutions, despite using different sets of explanation-seeking questions across the three studies and changing the items used to measure some of the potential triggers between Studies 2 and 3.

Factor scores were extracted from these factor models using Thomson's method (Thomson, 1951) and were used to predict ESC as in Study 2. For Study 3a, the factor scores for all three factors made a significant contribution to the prediction of ESC, learning: $\beta = 0.45$, 95% CI [0.41, 0.50], $\chi^2(1) = 341.18$, $p < 0.001$; export: $\beta = 0.25$, 95% CI [0.20, 0.29], $\chi^2(1) = 111.09$, $p < 0.001$; backward-looking: $\beta = 0.22$, 95% CI [0.17, 0.26], $\chi^2(1) = 91.67$, $p < 0.001$. Additionally, when a separate model was fit to predict ESC for each individual factor, the marginal R^2 of the learning model was 0.31, the marginal R^2 of the export model was 0.08, and the marginal R^2 of the backward-looking model was 0.11.

For Study 3b, the factor scores for all three factors made a significant contribution to the prediction of ESC, learning: $\beta = 0.30$, 95% CI [0.26, 0.34], $\chi^2(1) = 201.52$, $p < 0.001$; export: $\beta = 0.19$, 95% CI [0.15, 0.23], $\chi^2(1) = 77.89$, $p < 0.001$; backward-looking: $\beta = 0.37$, 95% CI [0.33, 0.41], $\chi^2(1) = 286.61$, $p < 0.001$. Additionally, when a separate model was fit to predict ESC for each individual factor, the marginal R^2 of the learning model was 0.19, the marginal R^2 of the export model was 0.06, and the marginal R^2 of the backward-looking model was 0.23. As predicted, backward-looking measures played a larger role in predicting ESC when questions were posed in response to a novel image stimulus. Nonetheless, expectations about learning and export still predicted ESC controlling for backward-looking considerations.

4.2.3. Factor structure and FSC

Next, we investigated whether the potential trigger measures in response to FSQs followed a qualitatively different factor structure from that for ESQs. To do so, we conducted a second set of exploratory factor analyses, using only the subset of the data in

[13] As in Studies 1 and 2, we tested the quadratic effect of information gap on curiosity. This revealed a significant quadratic effect in both Study 3a, $\chi^2(1) = 35.15$, $p < 0.001$, and Study 3b, $\chi^2(1) = 5.74$, $p = 0.02$. The magnitude of this effect was again fairly small in both studies, Study 3a: $\beta = -0.16$, 95% CI [−0.22, −0.11]; Study 3b: $\beta = -0.07$, 95% CI [−0.12, −0.01]. We include only the linear effect of information gap in all subsequent analyses.

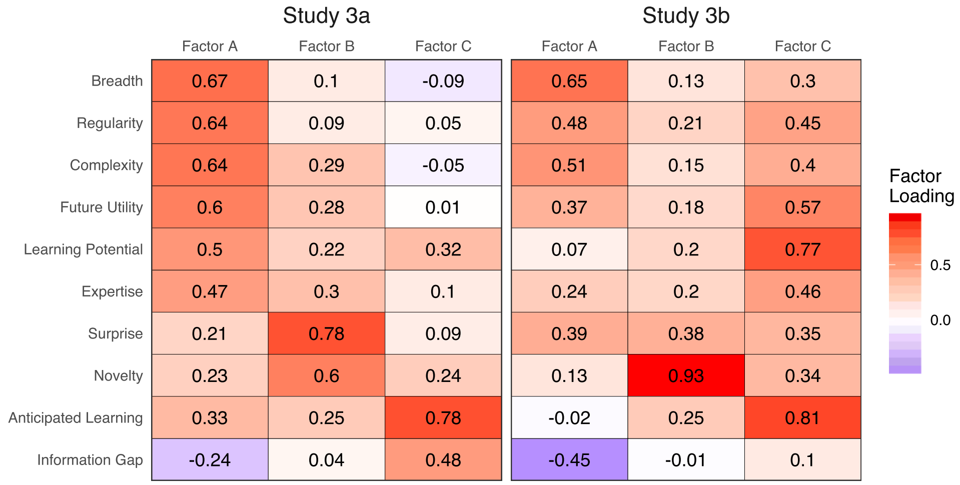

Fig. 3. Results of exploratory factor analysis on ratings in response to fact-seeking questions. Factors for Study 3b are rearranged to facilitate comparison with Study 3a (original order: factor C, factor A, factor B).

response to FSQs. As before, raw ratings were first group-mean centered on the basis of question to control for the nested structure of the data. For Study 3a, eigenvalues for the first three factors were over one, and explained 32%, 13%, and 9% of the variance in ratings, respectively. As parallel analysis suggested a two-factor solution, we compared a two- and three-factor solution. The three-factor solution produced better fit to the data than the two-factor solution; $TLI_{3\text{-factor}}$: 0.89, $RMSEA_{3\text{-factor}}$: 0.07, 95% CI [0.06, 0.08]; $TLI_{2\text{-factor}}$: 0.83, $RMSEA_{2\text{-factor}}$: 0.09, 95% CI [0.08, 0.10]. Therefore, we retained a three-factor solution.

An initial three-factor solution, using varimax rotation, was inspected. Negative valence and map mismatch failed to load onto any factor with a factor loading above 0.4 or any several factors with loadings above 0.3, and thus were eliminated. Excluding these items, the final three-factor solution with varimax rotation explained 47% of the total variance in ratings. The factors resulting from this analysis look notably different from the factors that emerged in response to explanation-seeking questions (see Fig. 3). The lack of consistency between the results of these two factor analyses precludes a more quantitative comparison.

For Study 3b, we initially fit a confirmatory factor analysis model to the FSC data (reported in the supplementary material), using the factor structure discovered in the previous analysis for Study 3a. However, this model did not achieve good fit by any standard metrics, so we conducted a second exploratory factor analysis. Initial eigenvalues for the first three factors were over one, and explained 38%, 11%, and 10% of the variance in ratings, respectively. We retained a three-factor solution, as recommended by parallel analysis. As in Study 3a, negative valence and map mismatch failed to load onto any factor with a factor loading above 0.4 or any several factors with loadings above 0.3, and thus were eliminated. Excluding these items, the final three-factor solution with varimax rotation explained 52% of the total variance in ratings. As can be seen in Fig. 3, the factors resulting from this analysis look qualitatively different from the factors that emerge in response to explanation-seeking questions across all studies. However, they are significantly though modestly correlated with the factor loadings found in Study 3a for FSQs, $r = 0.48$, $t(28) = 2.87$, $p = 0.008$.

As preregistered, we used the resulting factor scores to predict FSC. Due to the limited interpretability of these factors (and consequently, the limited interpretability of these regression analyses) we report these analyses in the supplementary material.

4.3. Discussion

In relation to Studies 1 and 2, we largely replicate our previous findings in Study 3. The relative strength of the 12 potential trigger measures in predicting curiosity was similar to the relative strength of these measures in Studies 1 and 2 (see Fig. 1). Further, though confirmatory factor analysis resulted in a poorly fitting model, exploratory factor analysis revealed a three-factor structure that was qualitatively and quantitatively highly similar to the three-factor structure found in Study 2 (see Fig. 2), across two different sets of explanation-seeking questions. The failure of the confirmatory factor analysis model despite the striking similarity between exploratory factor analysis solutions is likely at least in part driven by the constraint that each measure could only load onto one factor. As can be seen in Fig. 2, many measures had small but potentially meaningful factor loadings onto more than one factor. By limiting each measure to load onto a single factor, we limited the amount of variance this model could explain. Nonetheless, the exploratory factor analysis lends further support to the idea that ESC is driven by both backward-looking triggers and by forward-looking triggers that correspond to two functions—learning and export. We again found that triggers related to learning were highly predictive of ESC. Backward-looking triggers were much weaker in Study 3a, but stronger in Study 3b. This is potentially because the explanation-seeking questions in Study 3b were in response to novel image stimuli rather than familiar facts or phenomena—while the latter may be expected to elicit a fairly limited range of surprise or novelty, the former may be more sensitive to backward-looking considerations. In contrast to Study 2, however, triggers related to export were also somewhat weaker predictors of ESC across Studies 3a and 3b.

Study 3 additionally extends beyond Studies 1–2 by suggesting that ESC and FSC are elicited under somewhat different conditions. In Study 3a, anticipated learning was a stronger predictor of ESC than of FSC, while surprise at the content of a question was a stronger predictor of FSC than of ESC. In Study 3b, surprise and learning potential were stronger predictors of ESC than of FSC, while regularity and future utility were stronger predictors of FSC than of ESC. These were not the differences that we predicted (and, in the case of surprise, were inconsistent across studies), but they nonetheless suggest that future research might benefit from treating curiosity targeting explanations and curiosity targeting facts as potentially distinct. Supporting this suggestion, the latent variable structure of the tested triggers varied across the explanation-seeking and fact-seeking question sets. These differences leave open several avenues for future research on the distinction (if any) between ESC and other forms of curiosity.

5. General discussion

When and why do we ask why? Across three studies, we evaluated potential triggers of explanation-seeking curiosity: the state of curiosity about *why* something is the case that motivates a search for explanations. We proposed a functional approach to ESC premised on the idea that ESC should be triggered when pursuing explanations is likely to achieve core functions of explanation search: learning about the explanandum, learning for export, and learning from the domain-based expertise of others. Using this approach, we derived a set of potential triggers of ESC, including many that have not been tested in prior work on curiosity or information search. Notably, these potential triggers were forward-looking in nature—that is, concerning expectations about the downstream consequences of explanation search—in contrast to the backward-looking triggers (concerning fit with prior beliefs) that have dominated previous work on curiosity. Below, we review our empirical findings before turning to implications, open questions, and limitations.

First, we found that ratings of explanation-seeking curiosity successfully predicted choices about which explanations to reveal (Study 1). This validates our measure of explanation-seeking curiosity, as well as the link between ESC and explanation search that our functional approach assumes. Second, we found that forward-looking potential triggers predicted ESC above and beyond backward-looking triggers (Studies 1–3). These forward-looking triggers included anticipated learning, learning potential, need for relevant expertise, anticipated complexity, whether the explanandum is anticipated to reflect a pattern or regularity, and anticipated breadth of application. Notably, these included the three triggers with the highest predictive value across all of our data sets (learning potential, anticipated learning, and future utility).

To move beyond a large, unstructured list of individual predictors, we also explored the possibility that ratings of the potential triggers of ESC reflect a smaller number of latent factors (Studies 2–3). We identified three factors that were both highly interpretable and consistent across a diverse set of stimuli. These factors corresponded to (forward-looking) learning, (forward-looking) export, and backward-looking considerations. For fact-seeking questions, we found far less coherence and consistency (Study 3), which makes the coherence and consistency revealed for explanation-seeking questions all the more remarkable. Even for explanation-seeking questions, however, we did observe variation in the relative importance of each factor across studies, with the backward-looking factor playing the largest role when participants encountered novel explananda (Study 3b). These results simultaneously vindicate a role for backward-looking considerations from prior research (namely novelty or surprise) and suggest that forward-looking considerations (such as anticipated learning and future utility) deserve a great deal more attention.

Although our findings did support several candidate triggers of ESC drawn from past research, we found a minimal role for others. Our results do not support the role of negative valence in driving explanation seeking, as suggested in the literature on attribution (Weiner, 1985; Wong & Weiner, 1981), though this may be explained by the nature of our experimental stimuli, which did not include particularly valenced or personally-relevant questions. Our results also challenge the descriptive adequacy of theories of need for explanation proposed in philosophy (Grimm, 2008; W. Wong & Yudell, 2015)—across all studies, participant ratings of “map mismatch” and “fact-and-foil” were fairly weak predictors of ESC and of explanation search. This likely reflects both limitations in our measurement (a single rating item per theory) and limitations in the theories as descriptive accounts of human behavior.

An important question raised by our results is *how* people evaluate forward-looking considerations when encountering a potential target of explanation, given that these considerations are not reducible to the backward-looking triggers tested here. One reason backward-looking considerations may be imperfectly linked to forward-looking expectations is because the former are often inaccurate. Metacognitive calibration (i.e., how accurate a learner is in representing what they do and do not know) is often poor (e.g., Glenberg & Epstein, 1985; Lin, Moore, & Zabrucky, 2001). This has been shown for *explanatory* knowledge, in particular, with participants dramatically over-estimating their understanding of causal systems such as zippers and helicopters (Rozenblit & Keil, 2002). If people over-estimate how much they know, they are unlikely to experience the novelty, surprise, information gap, or map mismatch that would prompt curiosity, even in cases where they would benefit from acquiring further information. Notably, however, *attempting* to explain can improve metacognitive calibration: after Rozenblit and Keil (2002) prompted participants to generate an explanation, they produced lower (and more accurate) estimates of how well they actually understood the mechanism they were trying to explain. This suggests that better calibration (in particular, recognizing an information gap or map mismatch) may be a *functional consequence* of explanation search, and therefore an unreliable guide to whether explanation search is necessary in the first place. Similarly, some accounts of surprise suggest that surprise is the *consequence* of explanatory processes, not its cause: according to Foster & Keane (2015), surprise functions as a metacognitive signal reflecting the difficulty of explaining some observation (see also Maguire, Maguire, & Keane, 2011). These findings help explain why backward-looking considerations (such as information gap and surprise) may be so closely associated with the pursuit of explanations. However, we are left with the puzzle of how people form the judgements that actually trigger ESC, especially when these judgments concern expectations about the future. An important question for future research, then, is how people arrive at estimates of future learning or export, and whether such estimates are in fact reliable guides to how much will be learned from an explanation.

Relatedly, future research must go beyond our correlational approach to test whether forward-looking and backward-looking triggers can be experimentally manipulated to bring about ESC and explanation-seeking behavior. While forward- and backward-looking considerations will often go hand in hand, the relevant computations are conceptually distinct, and potentially dissociable empirically.[14] For example, the magnitude of one's surprise can be captured by the change in one's beliefs after receiving new evidence (Itti & Baldi, 2009), while anticipated learning can be captured by the expected reduction in uncertainty after asking a particular question (Oaksford & Chater, 1994). Given that the latter depends not only on one's current beliefs, but also on the particular answers one expects to receive, while the former depends only on current beliefs, these quantities may come apart under some circumstances (for example, a surprising piece of data may overturn a whole body of knowledge, while asking a single explanation-seeking question about that surprising data may only be expected to resolve a small amount of the resulting uncertainty, and only if particular answers are received). Future research is needed to determine the conditions under which backward- and forward-looking considerations do come apart, and how ESC depends on each set of considerations.

Another question for future research concerns the role of instrumental considerations in modulating the importance of various factors in driving explanation search. On many classic accounts of curiosity (e.g., Loewenstein, 1994), information search motivated by practical purposes (e.g., asking why a bicycle's brakes are squeaky in order to fix them) does not constitute curiosity at all. The present research, however, demonstrates that people's curiosity—at least as reflected in self-report—*is* in fact associated with expectations about future utility (see also Dubey & Griffiths, 2017, 2019). Importantly, neither our functional approach nor our data necessarily imply that participants *explicitly* evaluate export or utility prior to experiencing ESC, nor that they experience ESC as being purely instrumental. That said, we find it quite plausible that ESC for practical purposes (to the extent that this can still be considered curiosity at all) will be predicted more strongly by export-related considerations regarding future utility than will ESC for purely epistemic purposes.

Several additional limitations of our studies are worth acknowledging. First, while we used a large and diverse set of questions drawn from naturalistic sources, our stimuli and task were nonetheless restricted and somewhat artificial. With the exception of Study 3b, the questions that participants rated were largely within the domain of science, and with the exception of the ConceptNet question set, they induced a sufficiently high level of ESC or FSC to be posed in an online forum. It remains to be tested whether the triggers of ESC and explanation-seeking behavior are consistent across different sets of explanation-seeking questions (e.g., queries posted on online forums, queries entered into search engines, queries directed towards experts), across types of explanation-seeking questions (e.g., questions seeking teleological versus mechanistic explanations, or how- versus why- explanations), for different domains of questioning (e.g., questions concerning religious phenomena versus scientific phenomena versus social phenomena), in different cultural contexts, and when participants are required to generate questions themselves, rather than evaluating presented questions. Relatedly, while our focus was on the epistemic goals of explanation search, other research emphasizes the interpersonal or intrapersonal goals of explanation search (Malle, 2004) or the argumentative function of explanations (Mercier & Sperber, 2011). Future research might explore whether these functions suggest different potential triggers of explanation-seeking behavior.

Additionally, these studies leave open the questions of whether and to what extent ESC guides exploratory behavior more generally. For example, does ESC also predict when people seek information from experts, when people make or generate new observations, and when people attempt to generate explanations themselves? This final behavior is particularly important: some of the previous research we use to motivate the "functional" triggers of ESC are based on the functions of *internal* explanation search (i.e., explanation generation "in the mind"), while the behavior we use to validate our measure of ESC involves *external* explanation search (i.e., seeking the explanation from an outside source). It remains to be seen whether the same triggers predict "internal" versus "external" explanation search.

Broadly, this research provides a framework for thinking about the triggers of exploratory behavior, and of explanation search in particular. We find that explanation-seeking curiosity is aroused in a systematic and interpretable way. People are not only sensitive to the match or mismatch between a given stimulus and their current or former beliefs, but also to how they expect an explanation to affect their *future* selves. This future-directed orientation not only helps us understand curiosity in individuals, but also prompts new questions that can direct our future research as scientists.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank members of the Concepts and Cognition Lab for their useful comments. Some of the results reported here were presented at the 2018 meeting of the Society for Philosophy and Psychology and the 2018 meeting of the Cognitive Science Society, and we are grateful to these audiences for their discussion and feedback. This work was supported by a James S. McDonnell

[14] For example, in an unpublished study described in Loewenstein (1994), participants reported higher levels of curiosity about the rule that tied together a longer list of states versus a shorter list of states, purportedly because they wanted to learn the rule that was expected to unify *more* information (Loewenstein, Adler, Behrens, & Gillis, 1992). This manipulation seems to target forward-looking considerations with minimal influence on backward-looking considerations.

Foundation Scholar Award in Understanding Human Cognition and a Templeton Foundation Grant awarded to TL, as well as an NSF Graduate Research Fellowship to EL [grant numbers DGE-1752814 and DGE-1656466]. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the McDonnell Foundation, the Templeton Foundation, or the National Science Foundation.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cogpsych.2020.101276.

References

Aleven, V. A. W. M. M., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based cognitive tutor. *Cognitive Science, 26*(2), 147–179. https://doi.org/10.1207/s15516709cog2602_1.

Anderson, C. A., Krull, D. S., & Weiner, B. (1996). Explanations: Processes and consequences. In E. T. Higgins & Kruglanski (Eds.), Social psychology: Handbook of basic principles (pp. 271–296).

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked-out steps. *Journal of Educational Psychology, 95*(4), 774–783. https://doi.org/10.1037/0022-0663.95.4.774.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01.

Berlyne, D. E. (1950). Novelty and curiosity as determinants of exploratory behavior. *British Journal of Psychology. General Section, 41*(1–2), 68–80. https://doi.org/10.1111/j.2044-8295.1950.tb00262.x.

Berlyne, D. E. (1966). Curiosity and exploration. *Science, 153*(3731), 25–33. https://doi.org/10.1126/science.153.3731.25.

Berlyne, D. E., & Frommer, F. D. (1966). Some determinants of the incidence and content of children's questions. *Child Development, 37*(1), 177–189. https://doi.org/10.2307/1126438.

Blanchard, T., Lombrozo, T., & Nichols, S. (2018). Bayesian Occam's razor is a razor of the people. *Cognitive Science, 42*(4), 1345–1359. https://doi.org/10.1111/cogs.12573.

Blanchard, T., Vasilyeva, N., & Lombrozo, T. (2018). Stability, breadth and guidance. *Philosophical Studies, 175*(9), 2263–2283. https://doi.org/10.1007/s11098-017-0958-6.

Bonawitz, E. B., & Lombrozo, T. (2012). Occam's rattle: Children's use of simplicity and probability to constrain inference. *Developmental Psychology, 48*(4), 1156–1164. https://doi.org/10.1037/a0026471.

Bramley, N. R., Lagnado, D. A., & Speekenbrink, M. (2015). Conservative forgetful scholars: How people learn causal structure through sequences of interventions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(3), 708–731. https://doi.org/10.1037/xlm0000061.

Chi, M. T., De Leeuw, N., Chiu, M.-H., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*(3), 439–477. https://doi.org/10.1016/0364-0213(94)90016-7.

Chouinard, M. M. (2007). Children's questions: A mechanism for cognitive development. *Monographs of the Society for Research in Child Development, 72*(1), 1–129.

Coenen, A., Rehder, B., & Gureckis, T. M. (2015). Strategies to intervene on causal systems are adaptively selected. *Cognitive Psychology, 79*, 102–133. https://doi.org/10.1016/j.cogpsych.2015.02.004.

Dubey, R., & Griffiths, T. (2019). Reconciling novelty and complexity through a rational analysis of curiosity [Preprint]. https://doi.org/10.31234/osf.io/wg5m6.

Dubey, R., & Griffiths, T. L. (2017). A rational analysis of curiosity. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. J. Davelaar (Eds.). *Proceedings of the 39th annual conference of the cognitive science society* (pp. 307–312). Austin, TX: Cognitive Science Society.

Fitneva, S. A., Lam, N. H., & Dunfield, K. A. (2013). The development of children's information gathering: To look or to ask? *Developmental Psychology, 49*(3), 533–542. https://doi.org/10.1037/a0031326.

Flora, D. B., & Flake, J. K. (2017). The purpose and practice of exploratory and confirmatory factor analysis in psychological research: Decisions for scale development and validation. *Canadian Journal of Behavioural Science; Ottawa, 49*(2), 78–88. https://doi.org/10.1037/cbs0000069.

Foster, M. I., & Keane, M. T. (2015). Why some surprises are more surprising than others: Surprise as a metacognitive sense of explanatory difficulty. *Cognitive Psychology, 81*, 74–116. https://doi.org/10.1016/j.cogpsych.2015.08.004.

Frazier, B. N., Gelman, S. A., & Wellman, H. M. (2009). Preschoolers' search for explanatory information within adult-child conversation. *Child Development, 80*(6), 1592–1611. https://doi.org/10.1111/j.1467-8624.2009.01356.x.

Glenberg, A. M., & Epstein, W. (1985). Calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*(4), 702–718. https://doi.org/10.1037/0278-7393.11.1-4.702.

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. https://doi.org/10.1016/j.tics.2013.09.001.

Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution, 7*(4), 493–498. https://doi.org/10.1111/2041-210X.12504.

Grimm, S. R. (2008). Explanatory inquiry and the need for explanation. *The British Journal for the Philosophy of Science, 59*(3), 481–497. https://doi.org/10.1093/bjps/axn021.

Gruber, M. J., Gelman, B. D., & Ranganath, C. (2014). States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron, 84*(2), 486–496. https://doi.org/10.1016/j.neuron.2014.08.060.

Harris, P. L. (2012). *Trusting what you're told: How children learn from others.* Harvard University Press.

Harris, P. L., & Koenig, M. A. (2006). Trust in testimony: How children learn about science and religion. *Child Development, 77*(3), 505–524. https://doi.org/10.1111/j.1467-8624.2006.00886.x.

Hickling, A. K., & Wellman, H. M. (2001). The emergence of children's causal explanations and theories: Evidence from everyday conversation. *Developmental Psychology, 37*(5), 668. https://doi.org/10.1037/0012-1649.37.5.668.

Huang, F. L., & Cornell, D. G. (2016). Using multilevel factor analysis with clustered data: Investigating the factor structure of the positive values scale. *Journal of Psychoeducational Assessment, 34*(1), 3–14. https://doi.org/10.1177/0734282915570278.

Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research, 49*(10), 1295–1306. https://doi.org/10.1016/j.visres.2008.09.007.

Jirout, J., & Klahr, D. (2012). Children's scientific curiosity: In search of an operational definition of an elusive concept. *Developmental Review, 32*(2), 125–160. https://doi.org/10.1016/j.dr.2012.04.002.

Johnson, S. G. B., Valenti, J. J., & Keil, F. C. (2019). Simplicity and complexity preferences in causal explanation: An opponent heuristic account. 101222 *Cognitive Psychology, 113*. https://doi.org/10.1016/j.cogpsych.2019.05.004.

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*(8), 963–973. https://doi.org/10.1111/j.1467-9280.2009.02402.x.

Keil, F. C. (2003). Folkscience: Coarse interpretations of a complex reality. *Trends in Cognitive Sciences, 7*(8), 368–373. https://doi.org/10.1016/S1364-6613(03)00158-X.

Keil, F. C., Stein, C., Webb, L., Billings, V. D., & Rozenblit, L. (2008). Discerning the division of cognitive labor: An emerging understanding of how knowledge is clustered in other minds. *Cognitive Science, 32*(2), 259–300. https://doi.org/10.1080/03640210701863339.

Kominsky, J. F., Zamm, A. P., & Keil, F. C. (2017). Knowing when help is needed: A developing sense of causal complexity. *Cognitive Science, 42*(2), 491–523. https://doi.org/10.1111/cogs.12509.

Kon, E., & Lombrozo, T. (in prep). Why explainers take exceptions to exceptions.

Lastovicka, J. L., & Thamodaran, K. (1991). Common factor score estimates in multiple regression problems. *Journal of Marketing Research,* 105–112. https://doi.org/10.2307/3172730.

Legare, C. H. (2012). Exploring explanation: Explaining inconsistent evidence informs exploratory, hypothesis-testing behavior in young children. *Child Development, 83*(1), 173–185. https://doi.org/10.1111/j.1467-8624.2011.01691.x.

Legare, C. H., Mills, C. M., Souza, A. L., Plummer, L. E., & Yasskin, R. (2013). The use of questions as problem-solving strategies during early childhood. *Journal of Experimental Child Psychology, 114*(1), 63–76. https://doi.org/10.1016/j.jecp.2012.07.002.

Lin, L.-M., Moore, D., & Zabrucky, K. M. (2001). An assessment of students' calibration of comprehension and calibration of performance using multiple measures. *Reading Psychology, 22*(2), 111–128. https://doi.org/10.1080/02702710119125.

Liquin, E. G., & Lombrozo, T. (2018). Determinants and consequences of the need for explanation. In T. T. Rogers, M. Rau, X. Zhu, & C. W. Kalish (Eds.). *Proceedings of the 40th annual conference of the cognitive science society* (pp. 696–701). Austin, TX: Cognitive Science Society.

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin, 116*(1), 75–98. https://doi.org/10.1037/0033-2909.116.1.75.

Loewenstein, G., Adler, D., Behrens, D., & Gillis, J. (1992). *Why Pandora opened the box: Curiosity as a desire for missing information. Unpublished Manuscript, Department of Social and Decision Sciences.* Pittsburgh, PA: Carnegie Mellon University.

Lombrozo, T. (2007). Simplicity and probability in causal explanation. *Cognitive Psychology, 55*(3), 232–257. https://doi.org/10.1016/j.cogpsych.2006.09.006.

Lombrozo, T. (2016). Explanatory preferences shape learning and inference. *Trends in Cognitive Sciences, 20*(10), 748–759. https://doi.org/10.1016/j.tics.2016.08.001.

Lombrozo, T., & Carey, S. (2006). Functional explanation and the function of explanation. *Cognition, 99*(2), 167–204. https://doi.org/10.1016/j.cognition.2004.12.009.

Lutz, D. J., & Keil, F. C. (2002). Early understanding of the division of cognitive labor. *Child Development, 73*(4), 1073–1084. https://doi.org/10.1111/1467-8624.00458.

Maguire, R., Maguire, P., & Keane, M. T. (2011). Making sense of surprise: An investigation of the factors influencing surprise judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(1), 176–186. https://doi.org/10.1037/a0021609.

Malle, B. F. (2004). *How the mind explains behavior. Folk explanation, meaning and social interaction.* Massachusetts: MIT-Press.

Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences, 34*(2), 57–74. https://doi.org/10.1017/S0140525X10000968.

Mills, C. M., Sands, K. R., Rowles, S. P., & Campbell, I. L. (2019). "I want to know more!": Children are sensitive to explanation quality when exploring new information. *Cognitive Science, 43*(1), https://doi.org/10.1111/cogs.12706.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4*(2), 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x.

Nyrup, R. (2015). How explanatory reasoning justifies pursuit: A Peircean view of IBE. *Philosophy of Science, 82*(5), 749–760. https://doi.org/10.1086/683262.

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review, 101*(4), 608–631. https://doi.org/10.1037/0033-295X.101.4.608.

Pacer, M., & Lombrozo, T. (2017). Ockham's razor cuts to the root: Simplicity in causal explanation. *Journal of Experimental Psychology: General, 146*(12), 1761–1780. https://doi.org/10.1037/xge0000318.

Read, S. J., & Marcus-Newhall, A. (1993). Explanatory coherence in social explanations: A parallel distributed processing account. *Journal of Personality and Social Psychology, 65*(3), 429–447. https://doi.org/10.1037/0022-3514.65.3.429.

Rothe, A., Lake, B. M., & Gureckis, T. M. (2018). Do people ask good questions? *Computational Brain & Behavior, 1*, 69–89. https://doi.org/10.1007/s42113-018-0005-5.

Rozenblit, L., & Keil, F. C. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science, 26*(5), 521–562. https://doi.org/10.1207/s15516709cog2605_1.

Ruggeri, A., & Lombrozo, T. (2015). Children adapt their questions to achieve efficient search. *Cognition, 143*, 203–216. https://doi.org/10.1016/j.cognition.2015.07.004.

Ruggeri, A., Lombrozo, T., Griffiths, T. L., & Xu, F. (2016). Sources of developmental change in the efficiency of information search. *Developmental Psychology, 52*(12), 2159–2173. https://doi.org/10.1037/dev0000240.

Ruggeri, A., Sim, Z. L., & Xu, F. (2017). "Why is Toma late to school again?" Preschoolers identify the most informative questions. *Developmental Psychology, 53*(9), 1620–1632. https://doi.org/10.1037/dev0000340.

Schwitzgebel, E. (1999). Children's theories and the drive to explain. *Science & Education, 8*(5), 457–488. https://doi.org/10.1023/A:1008633709447.

Singmann, H., & Kellen, D. (2019). An introduction to mixed models for experimental psychology. In D. H. Spieler & E. Schumacher (Eds.), New Methods in Cognitive Psychology. Psychology Press.

Speer, R., Chin, J., & Havasi, C. (2017). *Conceptnet 5.5: An open multilingual graph of general knowledge. Thirty-first AAAI conference on artificial intelligence*.

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E.-J., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science, 27*(3), 453–489. https://doi.org/10.1016/S0364-0213(03)00010-7.

Strickland, B., Silver, I., & Keil, F. C. (2017). The texture of causal construals: Domain-specific biases shape causal inferences from discourse. *Memory & Cognition, 45*(3), 442–455. https://doi.org/10.3758/s13421-016-0668-x.

Thomson, G. H. (1951). *The factorial analysis of human ability.* London University Press.

Vasilyeva, N., Blanchard, T., & Lombrozo, T. (2018). Stable causal relationships are better causal relationships. *Cognitive Science, 42*(4), 1265–1296. https://doi.org/10.1111/cogs.12605.

Walker, C. M., Bonawitz, E., & Lombrozo, T. (2017). Effects of explaining on children's preference for simpler hypotheses. *Psychonomic Bulletin & Review, 24*(5), 1538–1547. https://doi.org/10.3758/s13423-016-1144-0.

Walker, C. M., Lombrozo, T., Williams, J. J., Rafferty, A. N., & Gopnik, A. (2017). Explaining constrains causal learning in childhood. *Child Development, 88*(1), 229–246. https://doi.org/10.1111/cdev.12590.

Weiner, B. (1985). "Spontaneous" causal thinking. *Psychological Bulletin, 97*(1), 74–84. https://doi.org/10.1037/0033-2909.97.1.74.

Wilkenfeld, D. A., Plunkett, D., & Lombrozo, T. (2016). Depth and deference: When and why we attribute understanding. *Philosophical Studies, 173*(2), 373–393.

Williams, J. J., & Lombrozo, T. (2010). The role of explanation in discovery and generalization: Evidence from category learning. *Cognitive Science, 34*(5), 776–806. https://doi.org/10.1111/j.1551-6709.2010.01113.x.

Williams, J. J., Lombrozo, T., & Rehder, B. (2013). The hazards of explanation: Overgeneralization in the face of exceptions. *Journal of Experimental Psychology: General, 142*(4), 1006–1014. https://doi.org/10.1037/a0030996.

Wong, P. T., & Weiner, B. (1981). When people ask" why" questions, and the heuristics of attributional search. *Journal of Personality and Social Psychology, 40*(4), 650. https://doi.org/10.1037/0022-3514.40.4.650.

Wong, W., & Yudell, Z. (2015). A normative account of the need for explanation. *Synthese, 192*(9), 2863–2885. https://doi.org/10.1007/s11229-015-0690-8.

Zemla, J. C., Sloman, S., Bechlivanidis, C., & Lagnado, D. A. (2017). Evaluating everyday explanations. *Psychonomic Bulletin & Review, 24*(5), 1488–1500. https://doi.org/10.3758/s13423-017-1258-z.