University of Reading

A reward-learning framework of knowledge acquisition: an integrated account of curiosity, interest, and intrinsic-extrinsic rewards

Article

Accepted Version

Murayama, K. (2022) A reward-learning framework of knowledge acquisition: an integrated account of curiosity, interest, and intrinsic-extrinsic rewards. Psychological Review, 129 (1). pp. 175-198. ISSN 0033-295X doi: 10.1037/rev0000349 Available at https://centaur.reading.ac.uk/102242/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1037/rev0000349

Publisher: American Psychological Association

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

University of Reading

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

A reward-learning framework of knowledge acquisition: An integrated account of curiosity, interest, and intrinsic-extrinsic rewards.

Kou Murayama[1]

University of Tübingen

[1] Hector Research Institute, University of Tübingen, Europastraße 6, 72072 Tübingen, Germany

Accepted by *Psychological Review* on Nov 9, 2021

Correspondence should be addressed to Kou Murayama,. Hector Research Institute, University of Tübingen, Europastraße 6, 72072 Tübingen, Germany. Email: k.murayama@uni-tuebingen.de. This research was supported by JSPS KAKENHI (Grant Numbers 16H06406, 18H01102; 18K18696), the Leverhulme Trust (Grant Number RL-2016-030); Jacobs Foundation Research Fellowship; and the Alexander von Humboldt Foundation (the Alexander von Humboldt Professorship endowed by the German Federal Ministry of Education and Research). Part of the proposed ideas were presented and discussed in several conferences and workshops including International Conference on Motivation (2018, Aarhus), Exploring curiosity (2018, Amsterdam), and Neurocuriosity (2016, London). The current article elaborated and integrated these earlier ideas after discussion and comments. An initial version of the manuscript was posted at https://osf.io/zey4k/. The script for the simulation is posted at https://osf.io/rsxuv/.

Abstract

Recent years have seen a considerable surge of research on interest-based engagement, examining how and why people are engaged in activities without relying on extrinsic rewards. However, the field of inquiry has been somewhat segregated into three different research traditions which have been developed relatively independently --- research on curiosity, interest, and trait curiosity/interest. We identify “long-term development” as a critical factor that links different research traditions, and set out an integrative perspective called the reward-learning framework of knowledge acquisition. This framework takes on the basic premise of existing reward-learning models of information seeking: that knowledge acquisition serves as an inherent reward, which reinforces people’s information-seeking behavior through a reward-learning process. Critically, however, the framework reveals how the knowledge-acquisition process is sustained and boosted over a long period of time in real-life settings (i.e., self-boosting effect), allowing us to integrate the different research traditions within reward-learning models. The framework also characterizes the knowledge-acquisition process with three distinct features that are not present in the reward-learning process with extrinsic rewards --- (1) selectivity, (2) vulnerability, and (3) under-appreciation. Finally, we discuss implications of the proposed framework regarding the debate over the conceptualization of broad concepts, namely; curiosity, interest, and intrinsic-extrinsic rewards.

Keywords. Intrinsic rewards; extrinsic motivation; information sampling; intellect, exploration

The construct of interest constitutes a fundamental aspect of human functioning. It supports an enormous variety of intellectual behaviors, ranging from early learning in children to scientific discovery and innovation. The role of interest in human functioning is unique in that it motivates people even *in the absence of* explicit extrinsic rewards (or extrinsic incentives; e.g., money, foods, prizes, social recognition, etc.). Interest endows us with remarkable capacity to sustain our motivation for activities without being fuelled by extrinsic rewards (Deci & Ryan, 1985; Renninger & Hidi, 2016).

Initial scientific research on interest (and related concepts) dates back to the 1950s (e.g., Berlyne, 1954; White, 1959) and received continued attention particularly in the literature of intrinsic motivation (Deci & Ryan, 1985). Following these traditional lines of research, recent years have seen a considerable surge of research focusing on this topic across a variety of fields. Indeed, a number of perspectives that seek to understand interest have been proffered in different disciplines, including developmental psychology (Engel, 2011; Jirout & Klahr, 2012; Twomey & Westermann, 2018), personality and social psychology (Kashdan, 2004; Litman, 2005; Mussel, 2013; Silvia, 2006), educational psychology (Hidi & Renninger, 2006; Muis, Chevrier, & Singh, 2018), organizational/vocational psychology (Hoff, Briley, Wee, & Rounds, 2018; Su, 2019), and aging (Sakaki, Yagi, & Murayama, 2018). The recent increase in attention is further boosted by theories offered by the field of computational modelling and neuroscience (Baldassarre & Mirolli, 2013; Gottlieb, Lopes, & Oudeyer, 2016; Gruber & Ranganath, 2019; Kidd & Hayden, 2015; Oudeyer & Kaplan, 2009). These different perspectives on interest have enriched the literature and provided many important insights into our understanding of this specific phenomena in respective fields. At the same time, however, they have been developed relatively independently and few attempts have been made to provide a complete picture of interest in human functioning.

One big impediment for the theoretical integration is the curious segregation of two major lines of studies; research on “curiosity” and research on “interest”. Although there are no agreed-upon definitions of these concepts even among researchers (Grossnickle, 2014; Hidi & Renninger, 2019; Markey & Loewenstein, 2014; Schmitt & Lahroodi, 2008), people would naively agree that curiosity and interest are obviously similar concepts. A recent survey showed that more than 80%

of both lay people and experts acknowledged that these two concepts have some overlap (Donnellan et al., 2021). Nevertheless, research on curiosity and research on interest have formed somewhat distinct research traditions only with perfunctory communication between them, with researchers developing different ways to approach, operationalize, and empirically study the concepts. For example, while recent research on curiosity provides many quantitative formulations of information-seeking behaviour (e.g., Baranes & Oudeyer, 2013; Mirolli & Baldassarre, 2013; Pathak, Agrawal, Efros, & Darrell, 2018; Schmidhuber, 1991), research on interest is characterized by more distinct focuses on the qualitative (i.e. non-quantitative) theorization of long-term task engagement (e.g., Deci & Ryan, 1985; Hidi & Renninger, 2006; Sansone & Thoman, 2005). These two research traditions even provide seemingly contradictory views on how interest motivates behavior (Braver et al., 2014; Murayama, 2018): On the one hand, research on curiosity tends to assume that both behavior driven by interest and by extrinsic rewards are governed by a common reward-learning process, emphasizing the similarities between interest-based engagement and engagement based on receiving extrinsic rewards (Murayama, 2018)[1]. On the other hand, research on interest supposes, either implicitly or explicitly, that interest is something special, and independent of (or even in opposition to) motivation driven by extrinsic rewards (e.g., Deci & Ryan, 1985). As a result of these opposing theoretical views, researchers even make contradictory suggestions on how to facilitate interest in applied settings (Howard-Jones & Jay, 2016; Kohn, 1993).

To further complicate the matter, curiosity and interest are often characterized at two different levels --- state and trait (Grossnickle, 2016). In fact, there is yet another line of research that specifically focuses on stable traits related to curiosity and interest (e.g., Campbell & Borgen, 1999; Kashdan et al., 2018; Litman & Spielberger, 2003), which we shall refer as research on trait curiosity/interest. Research focusing on stable traits examines people's overall preference for activities which often generalize across time and situations, and aims to identify different "types" of stable trait that can explain individual differences in such preferences. These studies demonstrated the importance of stable traits related to curiosity/interest in relation to other fundamental dimensions of personality, establishing an important position within the literature of personality psychology (e.g.,

Mussel, 2013). However, it has been rarely discussed to what extent the stable traits related to curiosity/interest are linked to curiosity and interest examined in other fields (for exceptions; see Jach et al., 2021; Silvia, 2006).

The purpose of this article is to integrate these three research traditions, providing a more complete picture of how people can sustain task-engagement without relying on extrinsic rewards. Critically, we illustrate the integration by conceptually extending the current reward-learning models of information seeking, which have attracted considerable attention in the curiosity literature (Gottlieb & Oudeyer, 2018; Gruber et al., 2014; Kidd & Hayden, 2015; Sakaki et al., 2018). The proposed *reward-learning framework for knowledge acquisition* helps us understand how these different theoretical perspectives are related to each other, and clarifies that both curiosity and interest are integral parts of the entire reward-learning process. The proposed framework also helps connect the existing reward-learning models in the curiosity literature to abundant theoretical perspectives and empirical findings in the literature of interest and trait curiosity/ interest, which none of the existent theoretical models have explicitly addressed. Note that we do not aim to provide a specific computational extension of the existing reward-learning models. Such a computational extension is still a big challenge given the complexities of modelling the core component of the proposed framework: knowledge representation and the self. In that respect, we prefer using the term *framework* rather than *model* to denote the proposed picture, suggesting that our aim is not to refute existing theories of curiosity or interest. Our aim is rather to provide a broad conceptual reward-learning framework that clarifies the key parameters that connect the three research traditions, providing a basis for the future development of the quantitative account of interest-based engagement in general.

In the following, we first briefly review each of the research traditions --- research on curiosity, interest, and trait curiosity/interest[2]. We then compare the research traditions and identify existing gaps that have prevented their integration. We show that these research traditions are especially different in terms of the extent to which “long-term development” is considered. We then put forth the reward-learning framework of knowledge acquisition by conceptually extending the existing reward-learning models from curiosity research. Critically, the proposed framework addresses long-term development by highlighting a self-boosting effect of interest-

based engagement, enabling us to integrate the three research traditions in a coherent manner. Following this, we present unique predictions derived from the proposed framework and their corroborating empirical findings. This is done by focusing especially on how the proposed framework can address differences between interest-based engagement and engagement based on extrinsic rewards (selectivity, vulnerability, and under-appreciation), without compromising the existing similarities between them.

Research on Curiosity and Reward Learning Models of Information Seeking

Research on curiosity is generally concerned with psychological and computational mechanisms underlying information-seeking behavior, especially when information bears no implications for acquiring extrinsic rewards (Gottlieb & Oudeyer, 2018). Research has shown that humans (and animals) often prefer acquiring information even if the information is not instrumental to earn rewards (e.g., Bennett, Bode, Brydevall, Warren, & Murawski, 2016; Blanchard, Hayden, & Bromberg-Martin, 2014; Bromberg-Martin & Hikosaka, 2009; Brydevall, Bennett, Murawski, & Bode, 2018; FitzGibbon, Komiya, & Murayama, 2020; Lanzetta & Driscoll, 1966; Rodriguez Cabrero, Zhu, & Ludvig, 2019; Wang & Hayden, 2019; Wyckoff Jr, 1952; see also Grant, Kajii, & Polak, 1998).

There are two classical theories in research on curiosity that attempt to explain people's motivation to acquire non-instrumental information: Berlyne's (1960, 1971) arousal theories and Loewenstein (1994)'s information-gap theory. But a third, more recent account is that reward-learning models can provide a viable framework to explain such non-instrumental information-seeking behavior (for a review, see Oudeyer & Kaplan, 2009). In traditional reward-learning models (Figure 1a), extrinsic rewards (e.g., money, food) reinforce behavior. Specifically, the value of an action in a situation is strengthened when the action leads to the acquisition of (especially unexpected) extrinsic rewards. Valued action is more likely to be performed again in the same situation, and after repeated cycles of the rewarded process, reinforced action may be extended to different contexts (i.e., generalization). Behavioral studies provide considerable support for the basic idea and various versions of the model have been proposed to provide fine-grained explanations of human behavior (for reviews, see Berridge, 2000; Dickinson & Balleine, 2002). One particularly important

recent advance was the finding that extrinsic rewards activate the so-called reward network (e.g., the ventral striatum) in the brain, indicating that these brain areas are responsible for updating and representing the reward value in the reward-learning process (Delgado, 2007; O'Doherty, 2004; Rushworth, Mars, & Summerfield, 2009).

Reward-learning models that explain information-seeking behavior presume that the basic principles of reward learning apply to interest-based information seeking (Figure 1b). Specifically, these models posit that information has inherent rewarding value, which reinforces our information-seeking behavior even without relying on explicit extrinsic rewards. According to these models, information-seeking behavior starts with awareness of a gap in our knowledge[3]. Note that “knowledge” here is broadly defined, and can include learning about specific topics (e.g., learning about psychology), honing skills (e.g., playing basketball; this is especially related to the concept of flow; Csikzentmihalyi, 1990), or resolving specific instances of uncertainty (e.g., determining the identity of ambiguous figures). Although the type of focal knowledge differs considerably between the examples (e.g., verbal knowledge, procedural knowledge, and perceptual knowledge), we consider all examples to pertain to knowledge acquisition.

In our daily life, we are always surrounded by uncertain stimuli. We cannot pay attention to everything that we do not know, but sometimes a certain lack of knowledge (i.e., knowledge gap) is made salient either explicitly or implicitly[4]. This awareness of a knowledge gap is often facilitated by so-called collative variables proposed by Berlyne (1966), such as novelty, uncertainty, conflict, and complexity. Once a knowledge gap becomes salient, people compute the *expected* reward value of the missing information, and it is this expected reward-value of information that motivates information-seeking behavior. Sometimes sought-after information is valued because it helps the agent to acquire more extrinsic rewards, but often agents value information for its own sake (i.e., information has an inherent value by itself). When information seeking is successful and reveals information that people were looking for, people experience rewarding value (often experienced as positive emotional feelings) from the acquired information. This rewarding experience would increase the expected reward value of the information-seeking behavior in the future (i.e., reinforcement).

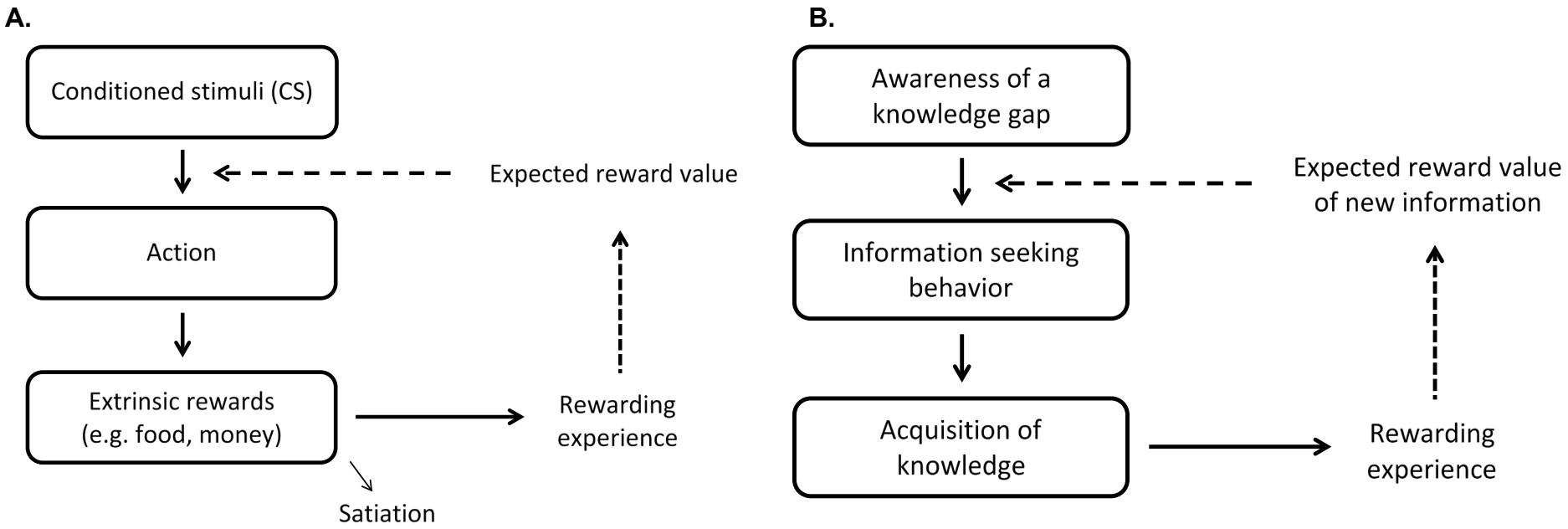


Figure 1 Reward learning models.

A. Standard reward-learning model with extrinsic rewards.

B. Standard reward-learning model that seeks to explain information-seeking behavior.

The idea that information-seeking behavior can be explained by reward-learning is not new. For example, Berlyne (1960) posited that the resolution of the state of novelty, uncertainty, conflict, and complexity (i.e., collative variables) is rewarding and this process further reinforces information-seeking behavior. However, recent work in the literature of curiosity advanced this idea in two important ways. First, recent studies have shown that knowledge acquisition activates the reward network in the brain like extrinsic rewards, suggesting that curiosity-based behavior and behavior based on extrinsic rewards are governed by a common reward-learning mechanism (Gruber & Ranganath, 2019; Fitzgibbon et al., 2020). For example, Jepma, Verdonschot, van Steenbergen, Rombouts, and Nieuwenhuis (2012) showed that the resolution of ambiguous pictures (i.e., reduction of perceptual uncertainty) increased activation in the striatum (see also Ligneul, Mermillod, & Morisseau, 2018; van Lieshout, Vandenbroucke, Müller, Cools, & de Lange, 2018, for more nuanced findings). Neuroimaging studies have also shown that the subjective feeling of curiosity in the face of knowledge gap (e.g., seeing a trivia question) activates the reward network in the same way that extrinsic rewards do (e.g., Gruber et al., 2014; Kang et al., 2009). Lau, Ozono, Kuratomi, Komiya, and Murayama (2020) directly compared the brain areas associated with the feeling of curiosity triggered by magic tricks and feeling of hunger induced by food pictures. They observed the shared activation in the striatum by these stimuli, which in turn predicted participants' behavior to satisfy their curiosity or hunger.

Second, research on cognitive science and computational modelling has produced numerous quantitative formulations of information-seeking behavior based on reward-learning models, with the aim to provide accurate understanding of exactly how and why people seek information. For example, Schmidhuber (1993) proposed a curiosity-driven reinforcement learning model in which agents are rewarded in proportion to the amount of learning progress one can expect when exerting an action. Review of these models is beyond the scope of the current manuscript (for reviews, see Baranes & Oudeyer, 2013; Mirolli & Baldassarre, 2013) but one of the most important areas of research in this field is to computationally formulate "what makes people curious in the first place", i.e., what types of information elicit more rewarding feelings than others. Different models highlight different aspects, such as uncertainty, learning progress, and perceived

competence/skills (for a review, see Oudeyer & Kaplan, 2009) and numerous models have been proposed to quantitatively formulate these aspects (e.g., Pathak et al., 2018; Shelhamer et al., 2017). These models are not always mutually exclusive, and curiosity is likely to be driven by a mixture of different reward-generating mechanisms (see de Abril & Kanai, 2018; Kobayashi, Ravaioli, Baranes, Woodford, & Gottlieb, 2019).

Research on Interest: Theories focusing on Development and Sustained Engagement

Compared to research on curiosity, research on interest is rather diverse. Different researchers propose relatively different theoretical perspectives, especially in the fields of educational psychology and social psychology. However, there is one distinctive aspect that many interest theories share: A focus on the long-term development of interest. Here we briefly review some of the major theories in the field by highlighting how long-term development is discussed as the core feature of interest.

Four-phase model of interest development

The four-phase model of interest development (Hidi & Renninger, 2006; Renninger & Hidi, 2016) is one of the most influential theories on interest, especially in the field of educational psychology. The theory includes four distinct phases where people develop interest. The first two phases are characterized by the term *situational interest*, and these represent task engagement that is largely dependent on available external resources such as learning materials and social relationships. The last two phases are characterized by the term *individual interest*, and represent relatively self-sustained task engagement, not contingent on availability of external resources. During these phases, people develop substantial knowledge and perceived value for the task, which in turn supports continued engagement over a prolonged period of time. The four-phase model of interest development underscores the importance of *valuing* knowledge for interest development (see also Wigfield et al., 2006), and this point has also received recent empirical support (Hulleman et al., 2017; O'Keefe & Linnenbrink-Garcia, 2014).

Person-object theory of interest

The person-object theory of interest is a broad theoretical framework developed by a group of researchers to explain voluntary task engagement (Krapp, 1999, 2000; Prenzel, 1992; Schiefele, 1991). The theory aims to explain how the selective persistence of task engagements --- the core characteristics of interest according to the authors --- develops over time. The theory supposes that interest lies in the interaction between persons (knowledge structure) and objects (environments), especially the cognitive conflict that arises as the gap between environmental input and one's pre-existing knowledge (or scheme). Importantly, the theory proposes that people's selective persistence manifests as people acquire new knowledge, because increased knowledge provides people with the ability or competence to resolve cognitive conflicts for specific task activities. The developed interest is considered to entail high subjective esteem (value) and positive emotional experiences for task activities with a rich amount of knowledge. In that respect, the theory critically distinguished interest from task engagement based on extrinsic rewards, which does not involve such characteristics (Krapp, 1999).

Emotion-attribution theory

In the emotion literature, interest has often been considered as a distinct emotion (e.g., Izard, 1977; Mortillaro, Mehu, & Scherer, 2011; Silvia, 2005; Smith & Ellsworth, 1988). In this area, Silvia has conducted the most extensive research program (e.g., Fayn, Silvia, MacCann, & Tiliopoulos, 2017; Silvia, 2006, 2008; Silvia, Henson, & Templin, 2009). Based on appraisal theories of emotion, Silvia argues that interest (or curiosity; he uses the term interchangeably) is an emotion caused by two components of subjective appraisal (Silvia, 2006). The first component is a *novelty check*, which judges whether the event is perceived as new or not. The second component is *coping potential (*or *comprehensibility)*, which refers to subjective estimates of the ability to understand and/or control the event. When an event is perceived as novel and comprehensible, the emotion of interest is generated (e.g., Turner Jr & Silvia, 2006).

Importantly, Silvia extended the previous idea, proposing a theory that focuses more specifically on the development of interest over the life span (Silvia, 2001; Silvia, 2006). Inspired by previous theories of interest or emotions (Allport, 1961; Prenzel, 1992; Tomkins, 1962), Silvia's emotion-attribution theory holds that

people's momentary emotional experience of interest and enjoyment form emotional knowledge --- subjective belief about the contingency between events and emotional experiences. Emotional knowledge is gradually shaped through repeated experiences and causal attribution processes, and it provides people with stable expectations of activities and events that are likely to produce experience of interest and enjoyment. Thus, the theory describes how repetitive momentary feelings of interest develop into a more stable form of interest (or interests).

Interest as intrinsic motivation

Intrinsically-motivated behavior refers to voluntary engagement in a task for the inherent interest and satisfaction derived from the task itself (Deci & Ryan, 1985). As this definition implies, interest and intrinsic motivation are closely related concepts (see also footnote 6 in General Discussion), but research on intrinsic motivation has followed somewhat different theoretical trajectories. Among them, the most influential theory is the self-determination theory, which argues that intrinsic motivation manifests as a consequence of the satisfaction of the three basic psychological needs: the need for competence, relatedness, and autonomy (Deci & Ryan, 1985; Ryan & Deci, 2017). According to this perspective, intrinsic motivation is a state in which people feel that they have control over the environment (competence), are supported by others (relatedness), and make their own choice (autonomy).

The theory has two critical features. First, in this line of research, the term "intrinsic motivation" is often used as a counter-concept of "extrinsic motivation". Extrinsic motivation refers to engagement in behavior for the purpose of obtaining extrinsic rewards. Importantly, in self-determination theory, extrinsic motivation has been conceptualized as potentially having detrimental effects on intrinsic motivation by thwarting the need for autonomy. One particularly popular phenomenon is the undermining effect, in which extrinsic rewards *decrease* people's voluntary engagement for a task when rewards are provided in a manner that impairs people's feeling of control (Deci, Koestner, & Ryan, 1999). The idea that intrinsic motivation is different from extrinsic motivation has become one of the predominant perspectives in the current field of motivation (Murayama, 2019), and has draws support from a

number of studies showing that intrinsic motivation and extrinsic motivation predict various different outcomes (e.g., Cerasoli et al., 2014).

Second, Ryan and Deci (2000, 2020) indicated that there is a continuum of different types of extrinsic motivation --- from behaviours regulated by externally imposed reward contingency (external regulation) to those regulated by one's own values and needs (integrated regulation). They argued that the satisfaction of basic psychological needs (autonomy, relatedness, and competence) helps people to internalize the value of the task, transforming the externally regulated activities into more autonomous task engagement. In other words, they highlight the importance of taking into account the potential long-term developmental process of task engagement. It is worth noting, however, that the theory still commits to a distinction between extrinsic and intrinsic motivation. Specifically, they hold that this internalization process occurs only for extrinsic motivation, and indicate that intrinsic motivation (i.e., behaviors driven by their inherent satisfaction) is separate and does not result from the transformation of extrinsic motivators. According to the self-determination theory, intrinsic motivation and internalized extrinsic motivation have a shared property (volitional engagement) but they are still distinct and not on the same continuum (Ryan and Deci, 2000, p 62; Ryan & Deci, 2017, p 197).

Research on trait curiosity/interest

In addition to research on curiosity and research on interest reviewed above, a third research tradition has examined curiosity and interest as personality traits (i.e., patterns of emotion, behaviour, and cognition that are relatively stable across time and situations). Although there are many models and theories that seek to understand the structure of trait curiosity and interest, they fall into two broad research traditions.

The first is an attempt to locate curiosity and interest within a comprehensive taxonomy of human personality based on extensive factor-analytic investigation. One example is Holland's (1985) model of vocational interest in the field of vocational psychology. Holland argues that individual differences in vocational interest consist of six dimensions that psychometrically form a hexagonal shape, covering all possible types of individual's belief and values: realistic, investigative, artistic, social, enterprising, conventional. In the field of personality psychology, curiosity and

interest are regarded as being one of the facets/aspects of Big-Five personality factors, especially the factor of openness to experience (Silvia & Christensen, 2020) and conscientiousness (von Stumm et al., 2011)

The second tradition aims to pinpoint theoretically meaningful subcomponents of trait curiosity and interest. For example, based on the wanting/liking distinction made in the neuroscientific literature (Berridge, 2000: Berridge & Robinson, 1998), Litman and Jimerson (2004; see also Litman, 2005) argue that there are two distinct personality traits related to curiosity. Specifically, they distinguish Interest-type (I-type) from Deprivation-type (D-type) trait curiosity. I-type trait curiosity reflects the tendency to experience positive emotional feelings induced by the anticipation of learning new information, whereas D-type curiosity reflects the tendency to experience a "feeling of deprivation", which triggers strong motivational intensity to resolve the uncertainty or knowledge gap. Kashdan et al. (2018) thoroughly reviewed the previous literature and developed a five-dimensional scale that includes the major aspects of trait curiosity in the literature: joyous exploration (i.e., I-type), deprivation sensitivity (i.e., D-type), stress tolerance, social curiosity, and thrill seeking.

Comparison of research on curiosity, interest, and traits

The literature review above highlights how the three lines of research on curiosity and interest differently approached the same phenomenon: motivational engagement without extrinsic rewards. We have summarized the key differences in Table 1.

Table 1 Comparison of the typical elements of the three research traditions

	Curiosity research	Interest research	Trait curiosity/interest research
Main fields	Cognitive science and neuroscience	Educational and social psychology	Personality and vocational psychology
Type of theories	Quantitative, mostly reward-learning models	Conceptual/descriptive, diverse	Conceptual/descriptive
Timeframe	Short-term change	Long-term change	Stable trait
Target phenomenon	Information seeking	Sustained engagement	Personality traits
Extrinsic rewards	Focus on similarities	Focus on differences	Little focus
Key focus	Uncertainty (collative variables), reinforcement	Development, valuation, knowledge structure	Individual difference
Main tasks/materials	Gambling tasks, trivia questions, visual images	Real-life learning materials, questionnaires	Questionnaires
Key question	What makes people curious?	How do people sustain long-term engagement?	What are the dimensions of curious people?

Research on curiosity is mainly concerned with people's motivation for information-seeking that does not necessarily bring future extrinsic rewards (Gottlieb & Oudeyer, 2018). Research on curiosity has become especially popular in recent years, with increasing numbers of researchers joining from the fields of cognitive science and neuroscience. Consequently, formal computational modelling is burgeoning and reward-learning models serve as one of the fundamental frameworks to study information-seeking behavior.

Reward-learning models suppose that knowledge acquisition serves as an inherent reward, and one of the critical questions in quantitative modelling is to understand exactly what types of information carry rewarding value and how agents detect knowledge gaps in the first place (Gottlieb et al., 2013; Oudeyer & Kaplan, 2009). In addition, as reward-learning models suppose that the common reward-learning system regulates behavior regardless of the reward type, researchers in the field of curiosity research tend to highlight the similarities between behavior driven by curiosity and that by extrinsic rewards. In empirical work, researchers have traditionally relied on experimental approaches, with the main focus on short-term (mostly one-shot) information-seeking behavior or associated momentary feelings. In fact, common experimental paradigms used to empirically examine curiosity are gambling tasks that provide participants with the opportunity to "peek" at the outcome of a gamble that cannot be changed (e.g., Blanchard et al., 2014; Bennett et al., 2016; Rodriguez Cabrero et al., 2019); or a series of textual or visual stimuli aimed to trigger momentary feelings of curiosity on a trial-by-trial basis, such as trivia questions (Kang et al., 2009), blurred images (Jepma et al., 2012), and magic tricks (Ozono et al., 2020).

Research on interest produced such a diversity in types of theories that some items in the summary provided in Table 1 may not apply to all of them, but they generally share a common theme: long-term development. Research on interest emerged from relatively applied fields (in comparison to curiosity research) such as educational and social psychology. For example, one of the early empirical examinations of people's interest was reading comprehension (e.g., Bernstein, 1955). As a result, the field is primarily interested in the developmental aspect, asking how and why people can sustain long-term commitments to learning and other activities without relying on extrinsic rewards. Theories of interest are largely

conceptual and descriptive, discussing how interest arises from an activity and changes into a more enduring form. The key factors identified in these theories are the personal valuation of activities and development of task specific knowledge (Deci & Ryan, 1985; Hidi & Renninger, 2006; Prenzel, 1992; Silvia, 2006). In addition, interest is often considered as overlapping with intrinsic motivation, and for that reason, interest has been considered as quite different, or even antagonistic to motivation driven by extrinsic rewards ("extrinsic motivation"). In empirical work, researchers prefer using real-life or educationally-relevant materials and topics, such as texts (Alexander & Jetton, 1996), school learning (Rotgans & Schmidt, 2011), and jobs (Gagné & Deci, 2005).

In the fields of personality and vocational psychology, curiosity and interest have been studied from a viewpoint of relatively stable traits --- the idea that there are individual differences in the extent to which people show curiosity or interest consistently over different situations or learning materials. This line of work is mainly concerned with finding possible personality dimensions of curiosity/interest. As such, the focus is on individual differences, and empirical studies mainly use self-report questionnaires administered to large number of participants, examining their factor structure and correlates. It is worth noting that, while many traditional theories simply describe the way people are classified (e.g., Holland, 1985), recent studies have taken further steps to inquire how personality traits of curiosity and interest are developed from a momentary feelings of curiosity (Fayn et al., 2019; Kashdan & Steger, 2007; Lydon-Staley et al., 2020). One perspective is that repeated momentary feelings are consolidated into more stable forms of curiosity and interest (see Fleeson, 2001). In that respect, research on interest and recent research on trait curiosity have an overlapping theme of inquiry.

The Key Question Linking the Three Research Traditions

Although this summary provides only a sketch of the diverse research conducted within the respective research traditions, the comparison clarifies two critical factors that distinguish these three lines of research: timescale and specificity of situations. While research on curiosity concerns how people are initially motivated to acquire knowledge in a specific situation, research on interest is more concerned with the long-term process of how individuals consolidate the initial motivation for

knowledge in real life. Research on trait curiosity/interest focuses on individual differences after people acquire stable interest for a particular broad topic (e.g., vocational interest) or general preferences for uncertainty (e.g., trait curiosity). Comparing research also highlights the strengths and weaknesses of existing reward-learning models of curiosity. While reward-learning models allow for accurate formulation of immediate information-seeking behavior in a specific situation by assuming that information gain is equivalent to the acquisition of extrinsic rewards, such a simplified assumption makes it difficult to explain the potential difference between interest-based engagement and engagement based on extrinsic rewards, which emerges over a longer period of time in real-life settings.

These observations lead us to a key question that needs to be addressed to close the gap between the three lines of research (see also Donnellan et al., 2021; Ten et al., 2021): How can we explain the long-term development of interest-based engagement within existing reward-learning models? Existing reward-learning models of curiosity have greatly advanced our understanding of temporal information-seeking behavior, but the models are not clear about how the temporal information-seeking behavior can be sustained and even bolstered over time (which is a focus of research on interest) or how it can be transformed into stable traits (which is a focus of research on trait curiosity/interest). Below, we introduce a framework to answer this question.

Reward-learning framework of knowledge acquisition

To fill gaps across the three different lines of research, we conceptually extend the existing reward-learning models of information seeking (Figure 1b) and propose a *reward-learning framework of knowledge acquisition* to fully represent the knowledge-acquisition process (Figure 2). As noted earlier, a critical feature of interest-based reward learning is that it relies on the self-generation of intangible rewards (i.e., knowledge) rather than the external supply of tangible rewards (i.e., food or money). We argue that it is this intangible and fluid nature of rewards that makes interest-based reward-learning critically distinctive from behavioral regulation based on extrinsic rewards, and it is this key feature that can explain the long-term development of interest-based engagement.

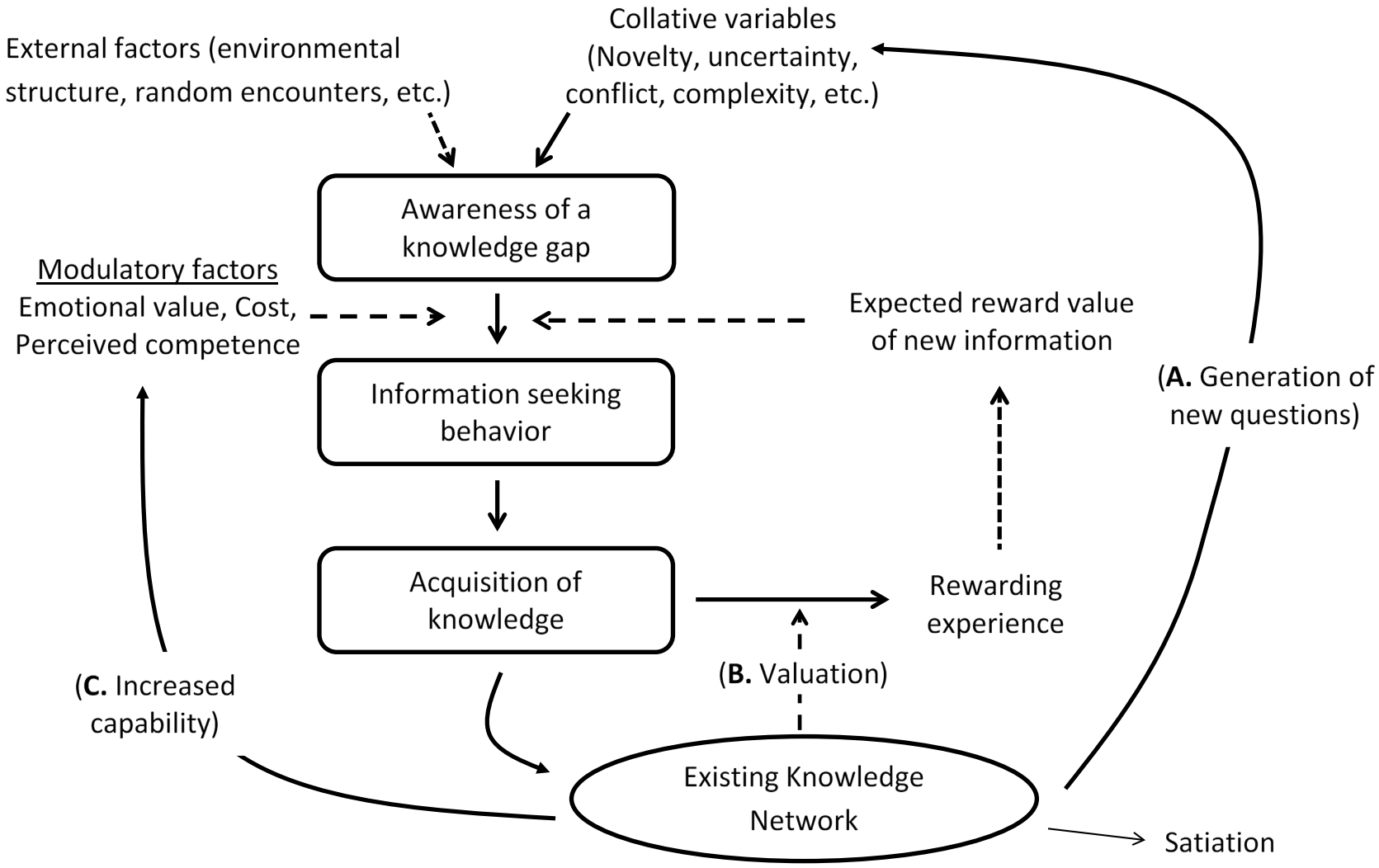


Figure 2. Proposed reward-learning framework of knowledge acquisition.

It is tempting to view knowledge acquisition as analogous to food acquisition. When we are deprived of food for an extended period of time, we feel hunger. Once we receive food, and consume it (with some excitement or disappointment, depending on the taste) we are then likely to become satiated. It is indeed easy to see the parallel between food and knowledge acquisition in this regard (Figure 1a and 1b). Consider a situation in which we come across trivia questions, which are standard materials used to induce feeling of interest in empirical work (e.g., Fastrich et al., 2017; Gruber et al., 2014). When we see a trivia question ("What food will never spoil?") we are likely to become interested in the answer. Just like with food, once we receive the answer to a trivia question, we normally appreciate the answer (with some excitement or disappointment, depending on our appraisal of the answer) and become satisfied.

In real-life learning materials, however, acquisition of knowledge is not an end point. Acquired knowledge is incorporated into a person's existing knowledge base and their updated knowledge base works in a way that facilitates further information-seeking behavior. There are at least three routes through which the updated knowledge base prompts further information-seeking behavior (Figure 2, Paths A-C). First, the acquisition of knowledge (e.g., learning multiple regression analysis in statistics) makes us aware of a "knowledge gap" that we were previously unaware of (e.g., we are now interested in how/whether standardized regression coefficients are different from previously learned concepts, e.g., partial correlation). This triggers further interest which in turn facilitates further motivation to acquire knowledge (Path A, "generation of new questions" in Figure 2). This is because enriching the knowledge network can highlight more room for expanding knowledge (Loewenstein, 1994; Renninger, 2000). Of course, the likelihood with which these processes occur depends on the nature of the task structure, a person's prior knowledge, and their skills (Prenzel, 1992). For example, learning the answer to a short, self-contained trivia questions (which is common in previous experimental studies on curiosity) may not be sufficient to trigger further interest in exploring the topic.

Second, it is also possible that accumulated knowledge makes us realize the importance of a topic (Path B, "valuation" in Figure 2), which further strengthens the reward value of new knowledge (e.g., we may learn how important it is to understand multiple regression analysis to complete our dissertation). This is because the value

of new information should depend on how much we can connect this new information to their existing knowledge. Finally, accumulated knowledge should also help us comprehend new information, which should modulate our decision to engage in information-seeking behavior. In other words, acquisition of new knowledge increases our perceived competence or skills (Path C, “increased capability” in Figure 2) required for further exploration[5].

Importantly, by taking into account the cumulative nature of knowledge, we can see that interest-based engagement is critically different from engagement based on extrinsic rewards. Specifically, knowledge acquisition and its incorporation into the existing knowledge base serves as a driver for future information-seeking behavior, making knowledge acquisition behavior sustainable. In other words, acquisition of new knowledge boosts motivation for acquiring further knowledge. Once this *self-boosting effect* kicks in, the reward system creates a positive feedback loop for knowledge acquisition, realizing long-term sustainability of interest-based engagement. This self-sustainability is in marked contrast with reward learning based on extrinsic rewards, in which sustainability hinges heavily on a constant external supply of extrinsic rewards. Indeed, self-sustainable engagement has been identified as one of the core characteristics of interest (re, e.g., Hulleman, Thoman, Dicke, & Harackiewicz, 2017; Renninger & Hidi, 2016; Sansone & Thoman, 2005; Silvia, 2008; Su, 2018). However, a critical point here is that the proposed framework explains the self-sustainability of interest-based engagement according to a reward-learning framework --- the development of interest is accounted for by the interaction between the reward-learning mechanism and inherent properties of knowledge acquisition. Thus, the proposed framework unifies the two research traditions (research on curiosity and research on interest) within a common scheme.

The reward-learning process can also underpin the formation of a stable trait supporting knowledge-acquisition behavior, thus linking it to research on trait curiosity/interest. Specifically, trait curiosity and interest can be considered a consequence of repeated cycle of reward-learning (i.e., reinforcement) and generalization (across time and situations) in the knowledge acquisition process. The idea that reward-learning underlies personality development has been discussed in the literature. For example, Wrzus and Roberts (2017), argue that repeated short-term sequences of reinforcement and generalization are the basic building block of

personality development (see also Back et al., 2011; Caspi & Roberts, 2001; Mischel & Shoda, 1995). Also, as noted earlier, recent research has theoretically posited that personality traits of curiosity and interest are developed from a repetitive experiences of the momentary feelings of curiosity (Fayn et al., 2019; Kashdan & Steger, 2007; Lydon-Staley et al., 2020). However, the proposed framework extends this idea through its self-boosting property and sustainability, suggesting that knowledge-acquisition process is inherently more likely than other psychological processes to manifest as a stable trait. In fact, the proposed framework can explain *why* there are so many identified personality traits which are related to curiosity and interest. Researchers have proposed a number of different dimensions of trait curiosity and interest (e.g., Kashdan et al., 2018; Litman, 2008), and there are also a number of other personality traits that have overlapping properties with curiosity and interest (e.g., novelty seeking, need for cognition, openness to experience, conscientiousness, etc.). Why are there so many personality traits that are related to curiosity and interest? According to the proposed reward-learning framework, this is because the knowledge-acquisition process involves a spontaneous generation of intrinsic rewards, increasing the likelihood of cycles of reinforcement and generalization taking place.

In sum, the reward-learning framework of knowledge acquisition extends existing reward-learning models by incorporating the fact that knowledge serves not only as a reinforcer (reward) but also as a basis for further information-seeking behavior. This is because critical stages of the reward-learning process --- awareness of the knowledge gap (Path A in Figure 2), evaluation of the rewarding value of new knowledge (Path B in Figure 2), and perceived competence to comprehend the new information (Path C in Figure 2) --- are the functions of the acquired knowledge itself, creating a complex recursive system. This is different from a system based on extrinsic rewards, where rewards are normally treated as an external supply without such recursive properties. As a consequence, the proposed framework indicates that knowledge-acquisition process affords a positive feedback loop of rewards that are internally generated, supporting sustainable engagement without relying on extrinsic rewards.

A quantitative illustration

We described three different routes through which the self-boosting effect of interest-based engagement is supported (Paths A-C in Figure 2). Among the three routes, perhaps the most counterintuitive one is the positive relationship between newly acquired knowledge and a person's awareness of knowledge gaps (Path A in Figure 2) --- i.e., the more we learn about a topic, the more we feel that we don't know. To provide a simple quantitative illustration, let us think about the effect on the knowledge network that some learning material affords for a learner. Figure 3 represents knowledge networks at an early, middle, and late stage of learning (Figure 3a-3c). Nodes represent pieces of information presented in the materials and edges represent the relationship of these different pieces of information. Pieces of information that have been learned (i.e., acquired knowledge) are represented in black. Pieces of information that have not been learned (so have not become part of an individual's knowledge base) are represented in gray (i.e., "unlearned" information). Hereafter, the size of the complete network (i.e., the sum of learned and unlearned nodes) is denoted as N (in Figure 3a-3c, $N = 14$).

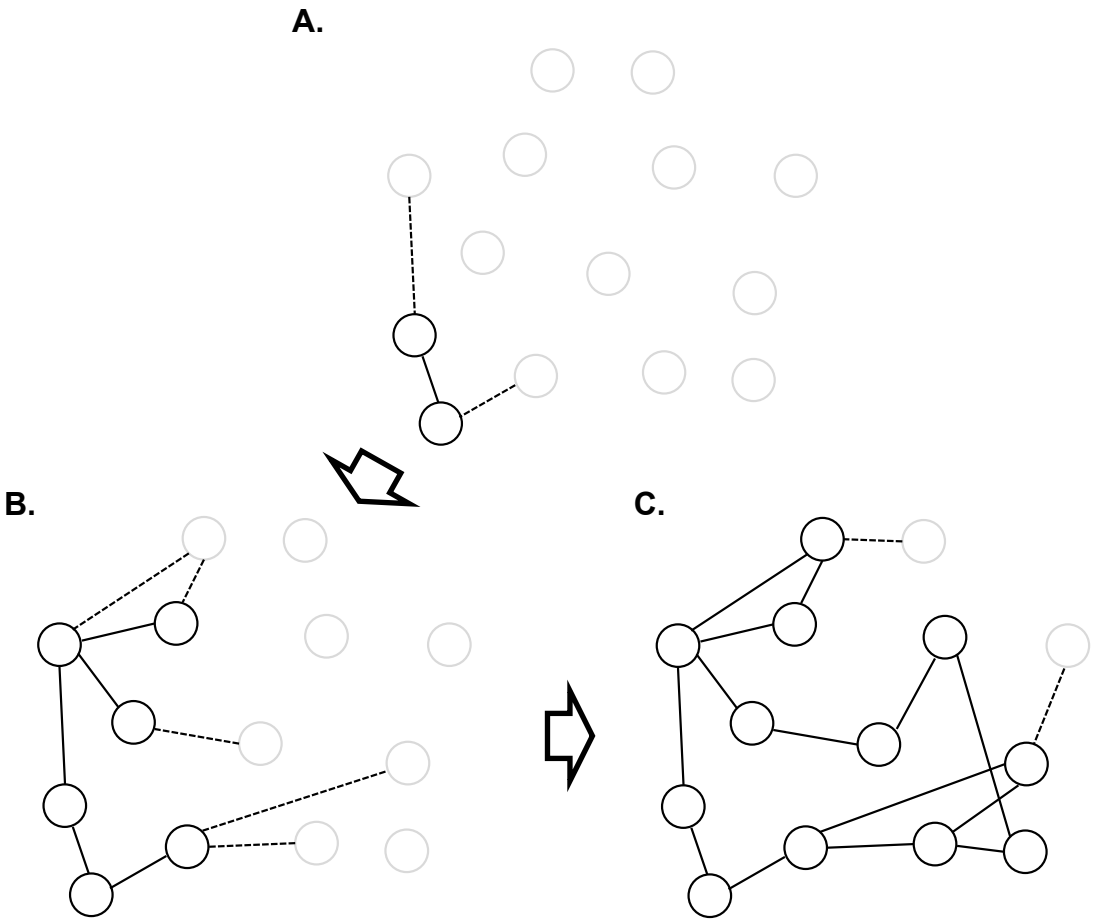
If we define a knowledge gap as the edge between acquired knowledge and unlearned information (dotted line in the figure), the number of knowledge gaps is initially small (Figure 3A) and increases as one learns information (i.e., acquires knowledge) (Figure 3B). The number of knowledge gaps eventually decreases as one approaches perfect mastery of the learning material (Figure 3C). This means that if one were to learn from material that consists of numerous pieces of information (i.e., N is very large), which is normally the case in real-life materials (e.g., learning physics), one should become aware of an increasing amount of knowledge gaps. Under the assumption that people randomly acquire knowledge from the network and that every piece of information in the learning material is randomly related to α% of the other pieces of information, learning the kth piece of information (i.e., acquiring the kth piece of knowledge) creates $\alpha(N - k)$ knowledge gaps in a person's knowledge base. This is the number of pieces of unlearned information multiplied by the probability that the unlearned information is linked to the kth piece of information (i.e., the kth piece of acquired knowledge). At the same time, it resolves $\alpha(k - 1)$ knowledge gaps (i.e., the new piece of information is expected to make $\alpha(k - 1)$ connections to nodes in their existing knowledge base). The expected number of knowledge gaps that *newly* arise as people acquire k pieces of

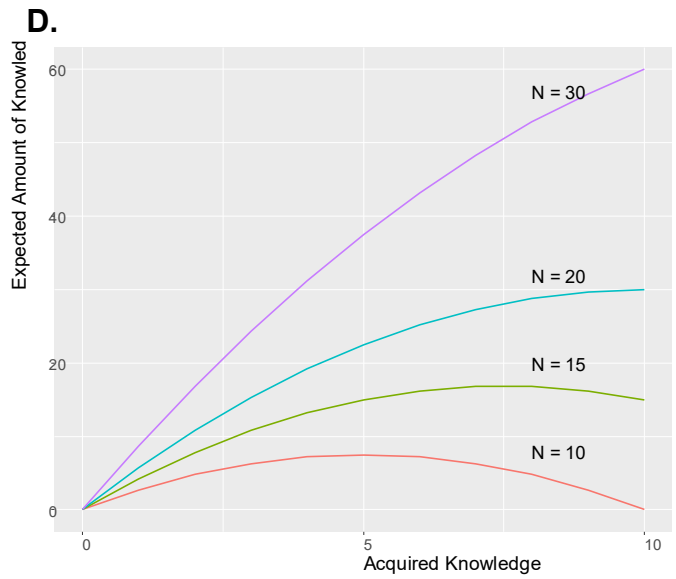
information can then be $\alpha(N - k) - \alpha(k - 1) = \alpha(N - 2k - 1)$. Consequently, the expected total number of knowledge gap when people acquire K pieces of information can be derived by summing the equation from $k = 1$ up to $k = K$: $\sum_{k=1}^{K} \alpha(N - 2k - 1)$. This results in the simple formula: $\alpha K (N - K)$. Using this equation, Figure 3D presents how the accumulation of knowledge (K) changes the expected amount of knowledge gaps in the network as a function of the network size N. Note that α in the equation simply serves as a scaling factor and therefore does not change the pattern of the curves. As can be seen, the number of knowledge gaps increases as people learn new information (acquire knowledge) as long the entire network size is large. Knowledge acquisition decreases the total amount of knowledge gaps *only* when one is learning a topic that comprises of a small, self-contained network, which is common with experimental materials in curiosity research.

Figure 3D assumes a random network as a representation of information, but previous studies have indicated that real-life knowledge networks are unlikely to be approximated by a random network (Steyyers & Tanenbaum, 2005). This is because knowledge networks tend to show small-world properties (e.g., Lydon-Staley et al., 2020; Tachimori et al., 2013). Accordingly, we simulated small-world networks (Watts & Strogatz, 1998) with various parameter values and a relatively large network size (100). We examined the robustness of the positive relationship between knowledge acquisition and the amount of knowledge gaps using these more realistic networks. As can be seen in Figure 3E, even with a small-world knowledge network, acquisition of knowledge generally increases the total amount of knowledge gaps when the amount of acquired knowledge is relatively small compared to the network size.

Of course, our illustration is a simplification. For example, the demonstration assumes that a person's knowledge network evolves at random, whereas in reality people purposefully learn new information (e.g., Wilson et al., 2014). Also, our definition of a knowledge gap is just one way of quantifying it. It is also worth noting that the existence of knowledge gap does not necessarily mean that one is aware of it (Loewenstein, 1994) --- people's awareness of knowledge gaps should be influenced by the recency of the knowledge acquired. Despite these potential complications, our simple illustration highlights one key fact: With little developed

knowledge, a knowledge gap is unlikely to be produced because there is no basis on which a “gap” is created, however “gap” is defined. In fact, Sizemore et al. (2018) quantified a knowledge gap as a void in the network and found a similar pattern on the relationship between knowledge acquisition and knowledge gaps, with various types of simulated networks and empirical knowledge networks. Christianson et al. (2020) also found a similar pattern using college-level mathematical linear-algebra texts to instantiate semantic networks.


A.
B.
C.


D.
Expected Amount of Knowled
60
40
20
0
N = 30
N = 20
N = 15
N = 10
0
5
10
Acquired Knowledge

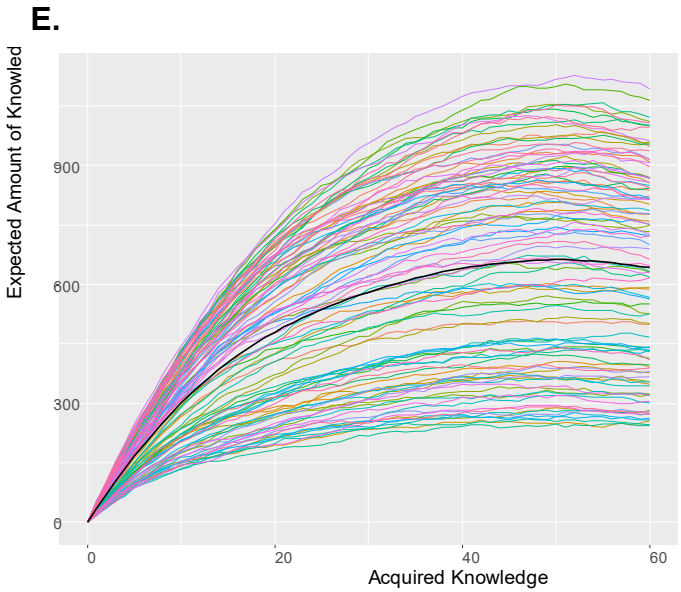

E.
Expected Amount of Knowled
900
600
300
0
0
20
40
60
Acquired Knowledge

Figure 3 Accumulation of knowledge and knowledge gaps. A-C. Knowledge network and knowledge gaps. Black nodes and solid edges represent acquired knowledge and gray-scaled nodes indicate unlearned information. Dashed edges represent knowledge gaps (information that is within the reach from the existing knowledge network but missing). As knowledge develops (A -> B), the amount of knowledge gaps increases. As knowledge approaches saturation (B -> C), the amount of knowledge gaps decreases. **D.** Expected amount of knowledge gaps as a function of acquire knowledge (x axis) and the knowledge network size (*N*). Unless *N* is small, knowledge acquisition generally increases the number of knowledge gaps in the network. The graph is drawn based on the equation in the main text assuming a random network, with $\alpha = 0.3$. **E.** Simulated 100 curves under the assumption of small-world network. Network size is fixed at $N = 100$. Black line represents the averaged curve. Small-world network was generated based on a ring graph with initial neighbourhood size = [10, 25] and rewiring probability = [0.1, 0.3] (randomly sampled from a uniform distribution for each curve). The script to reproduce the simulation/figure is available at https://osf.io/rsxuv/.

Empirical evidence

The self-boosting effect means that the effect of interest-based engagement becomes more noticeable over a longer period of time than engagement based on extrinsic incentives, because awareness of knowledge gaps and the rewarding value of new knowledge increases as one accumulates knowledge. Although we are not aware of the studies that directly tested the prediction, some empirical findings provide supportive evidence. For example, Murayama, Pekrun, Lichtenfeld, and vom Hofe (2013) revealed that students who exhibited high self-reported interest (intrinsic motivation) in mathematics in Grade 7 showed bigger increases in math achievement over 2 years (see also Gottfried, Marcoulides, Gottfried, & Oliver, 2013) than those who exhibited low self-reported interest. On the other hand, extrinsic motivation did not show such a long-term cumulative effect. The long-term cumulative effect of interest-based engagement and short-term effect of engagement based on extrinsic rewards have also been observed in both lab experiments (Murayama & Elliot, 2011), and field experiments (Vansteenkiste, Simons, Lens, Soenens, & Matos, 2005). Wrzesniewski et al. (2014) also showed that cadets (N > 10,000) were less likely to leave the military after the minimum required period if they were intrinsically motivated at the point of entry, but more likely to leave if they were extrinsically motivated (i.e., to make money).

Fastrich and Murayama (2020) provided more direct evidence that interest becomes stronger as one accumulates knowledge. In this study, university students viewed pieces of information one by one (e.g., demographics, role of women, language, etc.) about lesser-known countries (e.g., Burkina Faso) and reported their subjective interest about the country after the exposure to each piece of information. Participants could also stop viewing the information whenever they wanted (they could view up to 10 pieces of information). This paradigm made it possible to evaluate the piecemeal change in interest as a function of knowledge acquisition. Figure 4 shows the growth curve of interest estimated from data. Consistent with the proposed framework, when initial interest for a country was high (the top curve), participants reported increased feelings of interest towards the country as they accumulated knowledge about it, and continued information seeking until the end. Note that, towards the end of participants' information seeking, interest actually stopped increasing. This pattern of growth can be explained by the fact that the

learning material in the experiment is comprised of rather self-contained and isolated knowledge (see also Figure 3).

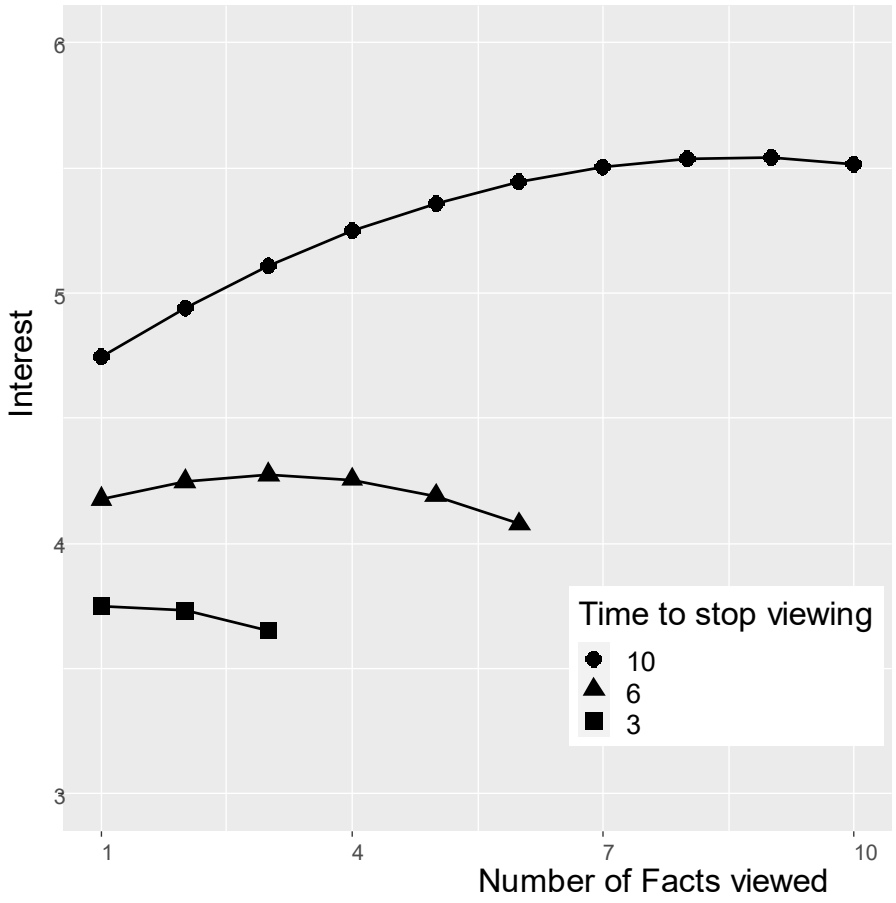


Figure 4 The relationship between the number of facts participants viewed and their interest, as a function of when they stopped viewing the facts. Figure was reconstructed based on Fastrich and Murayama (2020).

Other Predictions on The Difference between Interest-based Engagement and Engagement based on Extrinsic Rewards

We showed that the reward-learning framework of knowledge acquisition explains a critical feature of interest-based engagement: Self-boosting property. Below, we discuss three other predictions that the proposed framework makes regarding the differences between interest-based engagement and engagement based on extrinsic rewards: selectivity, vulnerability, and under-appreciation. On top of the self-boosting effect we discussed above, these features represent important characteristics that existing models of curiosity cannot readily account for. Existing models of interest and trait curiosity/interest may provide specific explanations for some of these features, but these explanations are often ambiguous about, or even deny the role of reward-learning mechanisms. Only the proposed framework provides a coherent account of these distinctive features of interest-based engagement, while simultaneously accommodating the common reward-learning processes between engagement based on interest and on extrinsic rewards.

Selectivity

The proposed framework suggests that the gap between those who are initially interested and those who are not should increase over time as people acquire more knowledge (i.e., the Matthew effect, snowball effect). People can sustain engagement for a long period of time once a positive feedback loop of knowledge acquisition is established, but in many cases people stop engaging before reaching this point. In our daily life, we become curious about many things that stimulate our knowledge gap, but this initial curiosity develops into long-term engagement only if there is a substantial initial commitment made for knowledge acquisition (Renninger & Hidi, 2016). In other words, the long-term effect of interest-based engagement works rather selectively for people who exhibit sufficient initial commitment.

The Matthew effect was observed in Fastrich and Murayama (2020) described above. In this study, although there was an overall increase in interest as participants acquired knowledge about lesser-known countries, this trend was smaller or non-existent for the countries for which participants did not have sufficient initial interest --- participants stopped seeking information earlier in these cases (the

middle and bottom curves in Figure 4). Accordingly, the initial gap in self-reported interest in a country widened as participants acquired more knowledge about the country. In fact, when participants exhibited low initial interest in a country, new knowledge about the country added little to their initial interest levels and participants were more likely to quit earlier (i.e., opt to switch to the next country). Similarly, Witherby and Carpenter (2021) showed in controlled experiments that participants' prior knowledge about football predicted their curiosity about football facts, and curiosity in turn predicted learning performance, producing a "rich-get-richer" effect. Critically, this effect was domain specific --- knowledge in other domains (cooking) was unrelated to curiosity in football. These results provide additional strong evidence for the selectivity of interest-based engagement.

These studies examined interest-based engagement for particular topics. However, the idea of selectivity more broadly predicts that there are considerable individual differences in specific topic(s) one is interested in: different people's interest manifests in a variety of different domains (e.g., math, music, etc.) and there is no single topic that everyone or no-one is interested in. This is because our initial engagement is influenced by so many idiosyncratic factors in our daily life. When an individual encounters a new topic, the topic may elicit some degree of initial feeling of interest in that individual, but this process would be sustained and bolstered only when the initial feeling is strong enough for a certain period of time to start the self-boosting cycle. Importantly, this initial engagement process (equivalent to situational interest described by Hidi & Renninger, 2006) is constrained by many internal (e.g., perceived ability, anxiety, prior knowledge and value), external (e.g., family and social support, and continued exposure to the topic), and even random factors (e.g., a book grabbed your attention when you were looking for a completely different book in a library). These factors may only have a small effect on initial engagement, but even a small difference in the initial engagement process may have long-term effects due to the self-boosting process. As a result, individuals with different initial internal and external conditions tend to develop different patterns of interest, and even individuals with apparently similar initial internal and external conditions can develop very different patterns of interest due to random and/or idiosyncratic factors as well as systematic factors.

The prediction is consistent with existing data on individual differences in curiosity and interest. For example, Mount, Barrick, Scullen, and Rounds (2005) conducted a meta-analysis on vocational interest and found that six types of vocational interest were only weakly or moderately correlated even after correcting for unreliability (average ρ = .28). Schoolchildren's self-reported interest in school subjects are also known to vary between individuals (see Bong, 2001; Goetz, Frenzel, Pekrun, Hall, & Lüdtke, 2007). Furthermore, by analyzing the responses to 244 trivia questions from nearly 2,000 participants, Fastrich et al. (2017) found that about half of the variance of self-reported interest ratings could be explained by the person x item interaction, meaning that individuals differed substantially in terms of which topics they found interesting. In fact, the data suggested that there is no single topic that was unanimously interesting (see also Ozono et al., 2020 for a replication of the findings using magic tricks as the material). Such large individual differences are not explicitly predicted by any other models of curiosity.

Vulnerability

Our framework supposes that extrinsic and intrinsic rewards generated by knowledge acquisition are both processed in the common reward-learning system, but that the knowledge-acquisition process enables continuous and sustainable generation of intrinsic rewards. In reality, our engagement is often supported by both extrinsic rewards and interest (e.g., professors teach not solely because they are paid to but also because they enjoy teaching). How do we regulate our engagement when both types of rewards are present?

In the reward-learning framework, this is essentially a problem of how you weigh each type of reward (i.e., how you define a function to combine multiple rewards to compute a single value score). In other words, this question can be framed as how humans *prioritize* different types of rewards. When we compare extrinsic rewards (e.g., money) and intrinsic rewards generated by knowledge acquisition, one obvious difference is salience: most extrinsic rewards are tangible and have clear incentive salience, whereas intrinsic rewards are invisible and their incentive salience is not immediately apparent (see also the section on "Under-appreciation" below). In addition, it is conceivable that the reward-learning system initially evolved to facilitate the acquisition of extrinsic rewards, which are directly

related to survival. As a result, the proposed reward-learning framework predicts that humans may prioritize processing extrinsic rewards when both extrinsic and intrinsic rewards are available. In other words, the presence of strong extrinsic rewards may prevent effective self-generation of intrinsic rewards through knowledge acquisition; interest-based engagement is sustainable, but at the same time it is vulnerable to the presence of salient rewards.

This prediction is supported by classic literature in social psychology: studies have shown that extrinsic rewards can *decrease* people's interest-based engagement for a task. For example (as explained earlier), studies have revealed the undermining effect, a phenomenon in which providing performance-contingent extrinsic rewards decreases people's voluntary engagement (or intrinsic motivation) for an inherently interesting task (also termed as "overjustification effect" or "motivation crowding out effect"; Frey & Jegen, 2001; Lepper et al., 1973). Previous meta-analyses on the undermining effect have suggested that extrinsic rewards can indeed decrease task engagement in certain circumstances (Deci, Koestner, & Ryan, 1999; Eisenberger, Pierce, & Cameron, 1999; Tang & Hall, 1995; Wiersma, 1992). Similarly, in classical work by Festinger and Carlsmith (1959), participants found a boring task more enjoyable when they were *not* sufficiently rewarded for doing the task ($1) compared to when they were highly rewarded ($20). This phenomenon is sometimes called the "insufficient justification effect" and demonstrates the possibility that people can self-generate rewards only when extrinsic rewards are not sufficiently available.

Critically, several studies have shown that the undermining effect is more pronounced when extrinsic rewards are salient (i.e. the existence of extrinsic rewards are perceived strongly by participants; Amabile, Hennessey, & Grossman, 1986; Eisenberger & Selbst, 1994; Hewett & Conway, 2016; Ross, 1975). These studies provide supportive evidence for the proposed reward-learning framework. It is worth noting that some theories that aimed to explain the undermining effect (Bénabou & Tirole, 2003; Deci & Ryan, 1985; Eisenberger et al., 1999; Zentall, 2005) also predict that extrinsic rewards are expected to undermine interest when they are made salient in such a way to detract people from the knowledge-acquisition process (i.e., when extrinsic rewards are preferentially processed). For example, a prominent account of the undermining effect comes from cognitive evaluation theory (Deci &

Ryan, 1985). According to this theory, extrinsic rewards undermine people's interest-based engagement when the "controlling aspect" of extrinsic rewards is made salient; that is, when people feel that they are doing the task just to obtain extrinsic rewards, rather than mastering and acquiring knowledge about the task. From the perspective of reward-learning framework, these theoretical explanations ultimately come down to the saliency of extrinsic rewards.

One critical difference between the reward-learning framework of knowledge acquisition and cognitive evaluation theory is that the proposed framework explains the undermining effect by extending existing reward-learning models. This fundamental assumption provides a unique prediction: activation in the reward-network produced by interest-based engagement should diminish after extrinsic rewards are initially provided and removed, because the initial provision of extrinsic rewards prevents the effective generation of intrinsic rewards in the reward network. This prediction has empirical support. For example, Murayama et al. (2010) conducted an experiment in which participants played with a game-like task with or without performance-based monetary rewards. They found activation in the striatum not only when participants received performance-based monetary rewards, but also when they were engaged in the task without any promise for such rewards. However, striatal activation was no longer observed when performance-based payment was provided and then removed, suggesting that extrinsic rewards prevent the self-generation of rewards (see also Ma, Jin, Meng, & Shen, 2014). Similarly, Izuma et al. (2010) showed that preference change caused by insufficient justification was associated with the change in the striatal activation in the reward network.

In addition to these empirical predictions, the idea that vulnerability of interest-based engagement can be explained by an extended reward-learning model has further theoretical implications. This is because undermining and insufficient justification effects are often cited as a phenomenon that cannot be fully explained by reward-learning models (e.g., Deci & Ryan, 1985; Linder, Cooper, & Jones, 1967). Specifically, researchers often argue that these effects cannot be explained by reward-learning models because provision of (sufficient) monetary rewards apparently decrease task motivation (when they are supposed to reinforce habit strength, e.g., Hull, 1943). However, these criticisms define rewards as specifically

extrinsic rewards (e.g., monetary rewards). If we define rewards more broadly, including intrinsic rewards generated by knowledge-acquisition process, then the undermining effect is not necessarily contradictory with our reward-learning framework.

Under-appreciation

One important, but often overlooked assumption in reward-learning models is that people have metacognitive capability to accurately estimate the expected reward-value upon receiving actual rewards. In other words, for reward learning to work effectively, people need to be able to accurately predict their affective or motivational states when they receive rewards. However, the literature on affective forecasting has indicated that people generally overestimate the affective value of future state --- called impact bias (Buehler & McFarland, 2001; Wilson, Wheatley, Meyers, Gilbert, & Axsom, 2000). There are not many studies that directly examine metacognitive accuracy relating to the effect of extrinsic rewards on affective and motivational states, but a few available studies have suggested that people tend to overestimate the affective and motivating power of extrinsic rewards, which is consistent with the idea of impact bias (Heath, 1999; Miller & Ratner, 1998; Nielsen, Knutson, & Carstensen, 2008; Sheldon, Gunz, Nichols, & Ferguson, 2010).

However, when it comes to the knowledge-acquisition process (Figure 2), metacognitive monitoring process seems to be much more complicated. In the case of extrinsic rewards (Figure 1a), rewards are salient, and we have an intuitive sense of how rewards feel (e.g., eating chocolate). In contrast, knowledge is intangible and often subtle; therefore it is easy to miss its potential affective and motivating values until one actually experiences it. It is also essential for people to metacognitively gauge the likelihood that they will successfully acquire knowledge (e.g., "tip of tongue"; Litman, Hutchins & Russon, 2005; Metcalfe, Schwartz, & Bloom, 2018). In addition, to make an accurate prediction about expected reward-value of knowledge acquisition, it is essential to take into account its self-boosting property. However, such a self-boosting process is something that people discover while learning, and it is difficult for people to take this factor into account before initiating a task.

Based on these observations the reward-learning framework of knowledge acquisition makes a novel prediction: as opposed to the impact bias observed for

extrinsic rewards, people *underestimate* the affective and motivating value of internal rewards generated by knowledge acquisition. A group of researchers have recently propose the idea of *metamotivation* or *metamotivational belief*, calling for research investigating the accuracy of our metacognitive ability to predict the consequences of different types of motivation (Miele & Scholer, 2018; Murayama, 2014; Scholer, Miele, Murayama, & Fujita, 2018).

Recent work in this field has provided evidence that people indeed underestimate the motivating value of interest. For example, in Woolley and Fishbach (2015), participants showed greater engagement in reading and rating interesting texts (jokes) in comparison to reading neutral texts (manual) but independent groups of participants failed to predict the positive effect of interesting text on task engagement. Ruan, Hsee, and Lu (2018) also found that people tend to avoid and underestimate their enjoyment of situations in which they are teased with missing information that is later delivered (see also Wilson, Centerbar, Kermer, & Gilbert, 2005). More direct evidence comes from Murayama, Kuratomi, Johnsen, Kitagami, and Hatano (2018; see also Hatano et al., 2020), in which researchers had participants work on a boring task (e.g., a 30min engagement in a flanker task without a break) and asked them to make a prediction about their degree of task engagement after doing the task for an extended period (with self-report questions or a behavioral index). This “predicted motivation” was compared with their “actual motivation” for the task, which was assessed during or after the task. Across multiple experiments, the results showed that people generally underestimated their actual motivation; participants predicted that the task would be quite boring, but they were actually engaged in the task more than they had expected (see “No Reward” condition in Figure 5). Interestingly, and consistent with the proposed reward-learning framework, this underestimation effect was not observed (or at least diminished) when they were provided with extrinsic rewards for engaging in the task (see “Reward” condition in Figure 5).

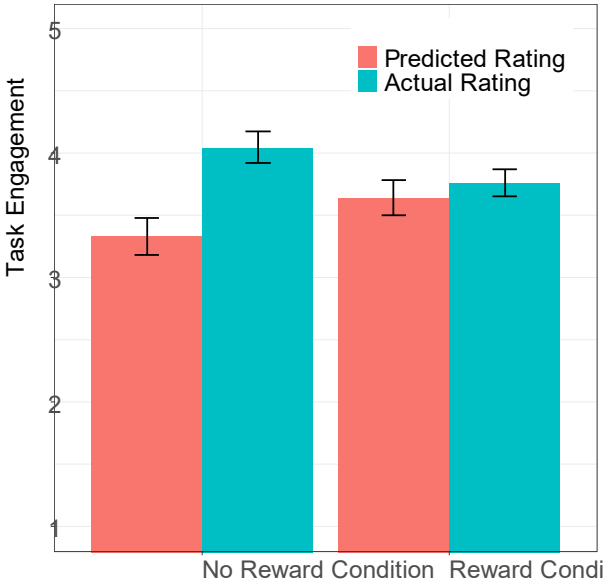


Figure 5. Comparison of predicted motivation and actual motivation as a function of reward with a flanker task. Participants underestimated task engagement when reward was not promised ("No reward condition") but this effect substantially diminished when reward was promised ("Reward condition"). Figure was reconstructed based on Murayama et al. (2018).

Previous studies have indicated that people tend to rely on extrinsic rewards when motivating others (Heath, 1999), even in situations in which extrinsic rewards are proven not to be empirically effective (Murayama, Kitagami, Tanaka, & Raw, 2017). Such a paradoxical phenomenon may be explained by metacognitive bias – that people tend to overestimate the effectiveness of extrinsic rewards and underappreciate their potential ability to sustain their engagement without extrinsic rewards. In previous literature on interest, metacognition has been examined as a correlate of interest (e.g., engaged learners tend to monitor and control their learning well; Pintrich & DeGroot, 1990; Tobias, 1995), and some studies discuss the importance of metacognitive monitoring in the psychological process underlying interest (Gruber & Ranganath, 2019; Litman et al., 2005; Loewenstein, 1994; Wade & Kidd, 2019). However, its accuracy (i.e., the extent to which people can accurately predict the rewarding feeling when they experience knowledge) has rarely been discussed. Understanding the metacognitive accuracy of interest (i.e., metamotivation of interest) may be another fruitful avenue for future research with great practical implications (see also MacGregor, 1960).

General Discussion

The research traditions of curiosity, interest, and trait curiosity/interest have each examined the knowledge acquisition process in the absence of extrinsic rewards. However, these three lines of research have developed independently, producing considerably different or even contradictory theoretical perspectives. We identified a key question to link the three research traditions: How can we explain the long-term development of interest-based engagement within existing reward-learning models? To answer this question, we proposed the reward-learning framework of knowledge acquisition as an extension of existing reward-learning models. The framework emphasizes the role of knowledge development in reward-learning processes, and explains long-term development by addressing the self-boosting property of interest-based engagement, thereby integrating the three different traditional lines of research.

Research on curiosity, interest, and trait curiosity/interest: How are they linked under the reward-learning framework?

Within the proposed framework, these three research traditions are considered to focus on different aspects of the same knowledge acquisition process. Curiosity research has mainly focused on short-term information-seeking behavior, aiming to answer the key question “What makes people curious?” (Table 1, left). This corresponds to the basic reward-learning process of information-seeking behavior in the proposed framework (i.e., the internal loop in Figure 2). Interest research, on the other hand, focuses on the long-term development of information-seeking behavior, aiming to answer the key question “How do people sustain long-term engagement?” (Table 1, middle). This corresponds to the elements that support the self-boosting effect in the reward-learning process (e.g., elements in the external loop in Figure 2). Research on trait curiosity/interest focuses on stable individual differences in relation to knowledge acquisition, aiming to answer the key question “What are the dimensions of curious people?” (Table 1, right). Such stable individual differences can be regarded as the consequence of the repetitive self-boosting processes in the proposed framework. Although there has been growing body of theoretical work on curiosity and interest (Dubey & Griffiths, 2020; Golman et al., 2020; Gruber & Ranganath, 2019; Kashdan et al., 2018; Sharot & Sunstein, 2020; Thoman et al., 2017), the proposed framework is the only account that integrates the three distinct research traditions from a coherent theoretical perspective. The proposed framework also clarifies commonalities as well as differences between behavioral regulation based on interest and on extrinsic rewards: While we maintained the common reward-learning properties between them, we identified several distinct features that represent interest-based engagement (i.e., self-boosting effect, selectivity, vulnerability, and under-appreciation).

On the distinction between extrinsic and intrinsic rewards

In this article, we tentatively distinguished extrinsic and intrinsic rewards. This dichotomy has been prevalent in the psychological literature, and the motivational processes driven by these rewards are often called extrinsic and intrinsic motivation respectively. However, the distinction between extrinsic and intrinsic motivation is much murkier than people suppose (e.g., Brief & Aldag, 1977; Dyer & Parker, 1975; Kruglanski, 1975). For example, social recognition (often called "social reward", e.g., Izuma, Saito, & Sadato, 2008) may not be extrinsically rewarding in that it is not a tangible reward, but it also may not be intrinsically rewarding as the rewarding

feeling does not come from the task or knowledge acquisition itself (Matyjek et al., 2020). As noted earlier, the self-determination theory also posits that extrinsic motivation is not a uniform construct and can be differentiated depending on the degree of internalization (Ryan and Deci, 2000, 2020).

Even if one is motivated for knowledge acquisition itself, it is not clear whether we can say that the person is purely intrinsically motivated. In fact, the reward-learning framework indicates that knowledge acquisition serves as an intrinsic reward, but this may be because knowledge generally helps people obtain extrinsic rewards (e.g., monetary rewards) in the long run. White (1959) argued that people are inherently motivated to understand and exert control over the external environment (i.e., competence) and this motivation for competence is seen as a critical aspect of intrinsic motivation (Deci & Ryan, 1985). However, if you consider *why* we are equipped with such motivation in the first place, a plausible answer is that such motivation would provide organisms with better access to extrinsic rewards in the external environment (i.e., intrinsic motivation has some adaptive value in evolution). In other words, it may be extremely difficult to find a completely pure form of intrinsic motivation (Kidd & Hayden, 2015).

This is also the case for extrinsic motivation. Consider a typical situation in which a person is rewarded with an extrinsic reward (e.g., money) for a particular action. It seems that the action is reinforced only by the extrinsic reward. However, upon scrutiny the situation is not that simple. That is, when the person is rewarded with the money, they also acquired the knowledge that the reward was contingent on the action. As a result, by understanding the contingency between the action and incentive, the person should experience a feeling of intrinsic reward (from knowledge acquisition; see also Ryan & Deci, 2020). This indicates that the action may actually be reinforced *both* by the extrinsic reward and the intrinsic reward produced by the knowledge acquisition. Consistent with this, studies show that the brain reward network is activated just by the knowledge of success (correct response), even if this knowledge does not come with any extrinsic rewards (Daniel & Pollmann, 2010; Tricomi, Delgado, McCandliss, McClelland, & Fiez, 2006).

One simple way of resolving the problem is to assert that we should not make any distinction among different types of reward. This idea has proven promising in

some empirical studies in neuroscience and computational modelling (e.g., Jones et al., 2011). However, this idea does not really explain why self-regulation based on intrinsic or extrinsic rewards involves so many different correlates, especially in real-world situations (Cerasoli, 2014; Kuvaas, Buch, Weibel, Dysvik, & Nerstad, 2017). Even in the neuroscientific literature, while studies show that extrinsic rewards and interest are processed in the common reward network (Lau et al., 2020; Murayama et al., 2010), they are also associated with unique brain activities (Domenico & Ryan, 2017; Reeve & Lee, 2019).

The reward-learning framework provides a more sensible solution. Specifically, the framework indicates that all types of rewards are processed in reward-learning mechanisms (i.e., they are processed on the same hardware) but how a specific type of reward works in the reward-learning process depends on its properties (i.e., different rewards are like different software). In other words, it is not the intrinsic-extrinsic dimension *per se* but specific and unique properties of rewards that matters. Pursuit of knowledge and pursuit of foods may share the same reward-learning mechanisms, but knowledge has unique properties that make the reward-learning process critically different from that with extrinsic rewards (e.g., self-boosting effect). Similarly, social recognition may serve as a reward to reinforce people's behavior but this does not mean that such social rewards work in exactly the same way as food or knowledge acquisition (see also Matyjek et al., 2020). Although there is a temptation in the literature to use the dichotomy (i.e., intrinsic rewards vs. extrinsic rewards) or singularity of rewards, we believe it is more important to think about the unique properties of different rewards on individual basis, articulating how such properties alter the way reward-learning works.

Are curiosity and interest distinct concepts?

The proposed framework puts us in a good position to consider a broader, but frequently-asked question --- what is the difference between curiosity and interest? Note that we are not asking the difference between the phenomena examined in curiosity research and interest research (which we already addressed in this manuscript), but the distinction of the concepts of curiosity and interest in general. We discussed this issue in detail elsewhere (Murayama et al., 2019), but in short, from our perspective, curiosity and interest are naïve concepts that people

subjectively construe from the knowledge-acquisition process. In fact, as can be seen in Figure 2, our framework does not include any elements called “curiosity” or “interest”, despite the fact that the framework aims to explain what people call curiosity and interest. In other words, both curiosity and interest may reflect post-hoc subjective labels (i.e. psychological construction; see also Russell, 2003) given to a myriad of phenomenological experiences which emerge during the knowledge-acquisition process. Because these concepts are subjectively constructed, not scientifically determined, we believe there exists neither strict definitions of curiosity and interest, nor a clear borderline between them[6]. This perspective also explains why different researchers define these concepts differently --- because curiosity and interest are the consequence of psychological construction, they allow every person to make their own subjective definition.

We are not claiming that curiosity and interest should be used interchangeably. What people call curiosity and interest may well be supported distinct subjective experiences and neural mechanisms (Donnellan et al., 2020). Some may want to call a specific part of the knowledge-acquisition process “curiosity” and another part of the process “interest” (this is what research on curiosity and interest implicitly does). Also, as reviewed in this manuscript, researchers in trait curiosity/interest have put forward various different "types" of curiosity, and these may also be supported by distinct parts of the reward learning process. In any case, as long as researchers share a common operationalization, this labelling certainly facilitates the scientific progress of curiosity and interest research. Our claim is simply that it is extremely difficult to determine the "right" definition of curiosity and interest, because there are no objective criteria to evaluate what is counted as curiosity or interest.

Related to the issue of terminologies, one of the challenges of curiosity and interest research is that the phenomena we normally refer to as “curiosity” and “interest” may be governed by multiple distinct mechanisms. The proposed reward-learning framework focuses on knowledge acquisition (and uncertainty reduction) but curiosity can arise for reasons other than knowledge acquisition. For example, recent studies have examined "morbid curiosity", which refers to people's motivation to see negative stimuli (Andersen et al., 2020; Andrade & Cohen, 2007; Hsee & Ruan, 2016; Oosterwijk, 2017; Oosterwijk, Snoek, Tekoppele, Engelbert, & Scholte, 2019).

While this type of morbid-seeking behavior can be explained by the motivation for knowledge acquisition or uncertainty reduction (i.e., people want to see something that they do not know, even if it is negative), it is also possible that people simply seek sensation itself (i.e., sensation-seeking; Zuckerman, 1979). When people explore the environment, we are tempted to assume that they are motivated (either implicitly or explicitly) to acquire information, but it is also possible that their exploratory behavior simply reflects the avoidance of the state of idleness or boredom (see Westgate & Wilson, 2018, for a comprehensive analysis of the state of boredom). Recent studies also showed that our curiosity is not only driven by the motivation for uncertainty reduction, but also by the motivation to savor the positive emotional valence of the outcome (Kobayashi, Ravaioli, Baranes, Woodford, & Gottlieb, 2019; Lieshout, Traast, de Lange, & Cools, 2019). Other researchers also indicate that people often seek information to reduce worry and gain emotional relief when they are confronted with threatening but potentially relieving information (Afifi & Weiner, 2006; Litman & Lunsford, 2010). When using the terms “curiosity” and “interest” in the scientific literature, we should be aware of the potential conflation of distinct psychological mechanisms underlying the phenomenon.

Factors that Inhibit Knowledge Acquisition

Our description of the framework is optimistic in that people are posited to always be committed to the knowledge-acquisition process because acquiring knowledge is rewarding. However, people are not always motivated to seek information. Indeed, research has shown that people decline to acquire information in some circumstances (Gigerenzer & Garcia-Retamero, 2017; Hertwig & Engel, 2016). Animal studies have also found animals’ tendency to choose novel stimuli (neophilia) or to avoid them (neophobia) depends on the circumstances (Hughes, 2007). In applied settings such as education, we know that students are not always motivated for learning (Hidi & Harackiewicz, 2000).

Some such instances, especially the deficit of interest observed in applied settings, may be well explained by the vulnerability of interest-based engagement explained earlier. However, explaining such phenomena in a more comprehensive manner requires incorporating some factors that modify the inherent positive reward value (Figure 2 “modulatory factors”; see also Sharot & Sunstein, 2020 for related

discussion). First, information-seeking behavior may be withheld when the behavior could potentially reveal negative information about the organism. For example, pregnant women do not always opt to take prenatal testing that informs them of the risk of genetic disorders (e.g., Down syndrome) for their expected child. In fact, Charpentier, Bromberg-Martin, and Sharot (2018) showed that participants' information-seeking behavior decreased when they expected to learn about bad outcomes. Second, the physical and mental cost of information-seeking behavior also serves as a barrier to initiating this behaviour. Indeed, in the studies described above, researchers found that organisms are often willing to seek information even if this involves some small cost (e.g., effort, monetary cost, time cost, etc.), but they also showed that increasing the amount of cost decreases the likelihood of information-seeking (e.g., Bennett et al., 2016; Blanchard et al., 2014).

Lastly, people should possess sufficient *perceived competence* (or *skills*) to execute information-seeking behavior. For example, even though we are interested in fully understanding Einstein's theory of general relativity, if we do not feel competent in physics and mathematics, it is unlikely that we will dare to start studying it. The role of perceived competence or skills[7] in the task engagement process is well documented in the literature of achievement motivation (e.g., Bandura, 1977; Deci & Ryan, 1985; Marsh & Shavelson, 1985; Wigfield & Eccles, 2002) and has been suggested as an appraisal component in the experience of interest (Silvia, 2006, which will be discussed later). However, the role of perceived competence has been relatively underexamined in the literature of neuroscience and computational modelling except for some notable exceptions (e.g., Barto et al., 2004; Schembri et al., 2007).

Note that although Figure 2 presents these factors as moderation variables, in reality, these factors operate in a more intricate manner in the knowledge-acquisition process. For example, when organisms acquire knowledge, that knowledge tends to also improve their subjective competence and skills. Therefore, knowledge and subjective competence co-develop over time and they may not be distinguishable in our learning system. In addition, because perceived skills are considered to be a form of knowledge, improving perceived competence or skills by itself should entail some rewarding experience (Mirolli & Baldassarre, 2013; Oudeyer & Kaplan, 2009). This is also consistent with the empirical literature that competence feedback for

skills is one of the important factors to facilitate interest (e.g., Harackiewicz, Manderlink, & Sansone, 1984). It is also possible that a cost itself acquires rewarding value if a costly investment (e.g., effort) returned valued information (i.e., learned industriousness; Eisenberger, 1992; see also Zentall, 2010). Although we do not discuss these inhibitory processes in more detail, we need to be aware of these complex dynamics when we fully model the knowledge-acquisition process.

Explanation of Other Concepts/factors Related to Curiosity and Interest

Previous studies have suggested that people's long-term engagement involves various factors that we have not accounted so far, such as goals and social contexts. Here we explain how these factors can be accommodated by the proposed framework. We do not intend to suggest that the reward-learning framework has better explanatory power than the existing theories that focus on these specific factors. Rather, our aim here is to show how we can understand these seemingly different theoretical constructs under a common mechanistic process using the reward-learning framework as a guiding scheme.

Goals

Goals play critical roles in interest-based engagement. In fact, previous studies have shown that the nature of goals influences the extent to which interest-based engagement is enhanced/impaired, such as whether there is congruence between means and the goal (Kruglanski et al., 2018), whether the goals are framed in terms of approach or avoidance (e.g., Elliot & Harackiewicz, 1996), whether the goals focus on inter- or intra-individual comparisons (performance vs. mastery goals; e.g., Dweck, 1986), whether the goal is to acquire knowledge itself or accomplishments/stimulation (Vellerand, 1997), and whether feedback type is aligned with goals (e.g., Sansone et al., 1989).

From the perspective of the proposed reward-learning framework, goals are part of an existing knowledge base, which modulates the rewarding value of upcoming information (Figure 2) --- e.g., if a person has a goal of outperforming others (i.e., performance goals), gaining knowledge from learning materials would not bring significant rewarding experience in comparison to a person who has a goal of task mastery itself (i.e. mastery goals; Dweck, 1986). The value of acquired skills

may be dependent on whether they serve to achieve the superordinate goal that one has (Kruglanski et al., 2018). This way of conceptualization is consistent with the idea that goals are cognitive representations (Elliot, 1999; Shah & Kruglanski, 2008) and part of self-knowledge (Conway & Pleydell-Pearce, 2000). It is also consistent with the idea that goals have a top-down influence on the valuation of activities (Austin & Vancouvder, 1996). Although the reward-learning framework of knowledge acquisition does not give fine-grained predictions about the effects of goals in comparison to the theories that focus on specific types of goals (e.g., Elliot & Harackiewicz, 1996), the framework highlights the importance of goals in information-seeking behavior and clarifies how goals should be studied and modelled in curiosity research.

Social context

Our interest does not always come from our own information-seeking behavior. We often become interested in a topic because we observe others (especially those who are close to us) being interested in the topic (Burgess, Riddell, Fancourt, & Murayama, 2018). For example, Dubey, Mehta, and Lombrozo (2021) showed that people's curiosity for knowledge questions was bolstered when they were informed that the questions were upvoted by other people. The reward-learning framework offers two potential mechanisms by which this so-called "social contagion of interest" takes place. First, observing others may simply make people aware of their knowledge gap, increasing their motivation to initiate information-seeking behavior. Second, it is also possible that observing the knowledge-acquisition process of others generates the rewarding feeling for the observer themselves, reinforcing the information-seeking behavior of the observer. This vicarious reinforcement process was originally proposed by Bandura (e.g., Bandura, Ross, & Ross, 1963) but empirical research has mostly been limited to situations in which participants observe others receiving extrinsic rewards (e.g., Burke, Tobler, Baddeley, & Schultz, 2010; Mobbs et al., 2009; Suzuki et al., 2012). Future studies should examine the role of vicarious reinforcement learning in explaining the influence of social factors on people's development of interest.

Flow

Flow experience refers to a mental state in which people are fully immersed in a particular activity, sustaining positive emotional feelings and engagement in an autotelic manner (Csikszentmihalyi, 1975). The literature suggests that people enter into a mental state of flow when there is a balance between skills and challenge, predicting an inverted-U relationship between skill level and flow state (Csikszentmihalyi, 1990). A person would not experience flow when their skill-level is too low (which leads to anxiety) or too high (which leads to boredom). We view flow as a strong phenomenological state underlying interest-based engagement. According to the reward-learning framework of knowledge acquisition (Figure 2), when the system forms a positive feedback loop and the self-boosting effect kicks in, the agent constantly experiences rewarding feelings that originate from knowledge acquisition, exhibiting sustained engagement in information-seeking behavior without relying on extrinsic incentives. This state is exactly what the flow experience is about.

Although speculative, this perspective can also explain the theoretically predicted inverted-U relationship between skill-level and flow state (Csikszentmihalyi, 1990). The skill factor in flow theory corresponds to perceived competence in the reward-learning framework (Figure 2). When perceived competence is very low, even if one expects a large knowledge gain, the person is unlikely to initiate information-seeking behavior, because the person is not confident that they will comprehend the new information. When perceived competence is very high, the person is confident that they will comprehend the information, but the expected knowledge gain is relatively small because they already know the material very well. As a result, people's information-seeking behavior is most sustained when skill level is perceived as moderate (see also Atkinson, 1957; Silvia, 2008, for a similar argument). Future studies would benefit from more rigorous quantification of this conceptual-level explanation with supporting empirical data.

Future directions

A few future directions should be noted. First, our proposed framework sheds new light on how we understand the developmental trajectories of curiosity and interest. There are common beliefs about the developmental changes of curiosity and interest, such as 'younger children are more curious' and 'older adults lose

curiosity and interest as they age' (see Aslan et al., 2021). Some research findings support this view: Previous studies using survey questions have found that curiosity and interest in school-aged children decreases over time (Gottfried, Fleming, & Gottfried, 2001; Spinath & Steinmayr, 2008; Wigfield et al., 2006), and that older adults exhibit decreased curiosity in comparison to younger adults (Robinson, Demetre, & Litman, 2017; Ziegler, Cengia, Mussel, & Gerstorf, 2015).

However, the reality is much more nuanced. In light of our reward-learning framework of knowledge acquisition, such a view risks missing the whole story about curiosity and interest. For example, while it may be true that younger children exhibit active and momentary information-seeking behavior more frequently than older children, apparent information-seeking behavior is just one aspect of knowledge-acquisition process. Younger children may not develop a sufficient knowledge base, and as a result, they may not be able to sustain information-seeking behavior over a long period of time (see also Frenzel, Pekrun, Dicke, & Goetz, 2012). It is also possible that older children are more engaged in mental information-seeking (i.e., sense making) processes rather than overt information seeking behavior; younger children may look more curious simply because their information-seeking behavior is more visible. In fact, Schultz, Wu, Ruggeri, and Bjorn (2019) found that young children were more likely than adults to explore their environment to reduce uncertainty, but young children failed to generalize the outcomes to similar (in this experiment, spatially proximal) situations (but see also Gopnik, Griffiths, & Lucas, 2015; Somerville et al., 2017). In a similar vein, because older adults tend to have a richer knowledge base than younger adults, there may be more room for older adults to develop long-term sustained engagement (Sakaki et al., 2018). It should also be noted that information-seeking behavior happens as a function of multiple factors, such as the rewarding value of knowledge, the cost of behavior, and perceived competence as well as environmental demands. Therefore, when a person actively seeks or does not seek information, there are multiple potential mechanisms underlying their decision. When interpreting observed age differences in information-seeking behavior, we need to be aware of these differential mechanisms underlying the behavior.

Second, on a practical note, the existence of a common reward-learning system raises an interesting possibility that extrinsic rewards may be utilized to

facilitate people's interest-based engagement (Murayama, 2018). There are mixed views on the utility of extrinsic rewards in education (Howard-Jones & Jay, 2016; Kohn, 1993) and our proposed framework may provide a potential compromise. Specifically, our framework indicates that people have a capacity to self-boost interest-based engagement over time for themselves, but this process can take place only when the knowledge-acquisition process has been started up. When the knowledge-acquisition process is difficult to initiate for certain reasons (e.g., people may not appreciate the potential value of knowledge in the first place) providing extrinsic rewards to engage people in the knowledge-acquisition process may not be a bad idea, as long as we ensure that people are not reliant on extrinsic rewards in the long run (in that case, the undermining effect happens). In other words, even when one was initially forced or incentivised to do a task, this initial engagement in the task can provide an opportunity for the person to have rewarding experiences from knowledge acquisition, potentially activating the self-boosting process. This "motivational transformation" from extrinsic to intrinsic rewards has been suggested by some theories (e.g., Hidi & Renninger, 2006), and there have been a few empirical studies supporting the idea. For example, Woolley and Fishbach (2016) showed that providing initial rewards (e.g., snacks) to high-school students helped their immediate rewarding experience, eventually increasing their enjoyment and persistence for school assignments (see also Harackiewicz, Manderlink, & Sansone, 1984). Future studies should examine more detailed psychological mechanisms underlying the motivational transformation process.

Finally, although we provided evidence that feelings of curiosity or interest can reinforce our information-seeking behavior, there are only a limited number of such studies and most focus on novelty in perceptual stimuli (e.g., Bevins, 2001; Duzel et al., 2010; Kish, 1955; Reed et al., 1996; Wittmann et al., 2008). In addition, these studies examined simple reinforcement of behavior, with little research investigating the exact properties of the feeling of curiosity or interest as a reinforcer, such as resistance to extinction and selective devaluation. Future studies should directly address the reinforcement properties of curiosity and interest in the knowledge-acquisition process.

Concluding comment

Knowledge acquisition is commonplace in our daily life. Every moment, we are acquiring new knowledge and updating our existing knowledge base. A great deal of research on cognitive psychology and neuroscience (e.g., memory, category learning) has examined how information is encoded, stored, and retrieved in our knowledge system. However, these studies have paid little attention to the potential motivating function of the knowledge-acquisition process itself (for exceptions, see Stojic, Schulz, Analytis, & Speekenbrink, 2019; Lydon-Staley, Zhou, Blevins, Zurn, & Bassett, 2019; Twomey & Westermann, 2018). On the other hand, research on curiosity and interest has investigated how knowledge-acquisition behavior is motivated, but these studies often do not consider how our knowledge is stored and updated in our memory and belief system (see also Zurn & Bassett, 2018). Our reward-learning framework of knowledge acquisition emphasizes the importance of connecting these two research fields in order to achieve a comprehensive understanding of humans' ability to sustain task engagement in the absence of extrinsic rewards.

Footnotes

1. In this article, we use the term “engagement” simply to refer to commitment to or involvement with a task. “Engagement based on extrinsic incentives” refers to task commitment primarily driven by extrinsic rewards, whereas “interest-based engagement” refers to task commitment based on internal rewarding feeling arising from knowledge acquisition (which will be explained later).

2. In the current article, we use the term “curiosity” and “interest” in a manner that lay persons use them, and do not provide strict definitions. This is because we view the terms “curiosity” and “interest” are subjective construction of reward-learning process in knowledge acquisition, and therefore it is inherently very difficult to precisely define/distinguish them. We will discuss the issue of definitions in General Discussion (see also Murayama et al., 2019).

3. In the field of information seeking and sampling, researchers use the term information rather than knowledge to describe the knowledge-acquisition process that we discuss in this paper. We use the terms interchangeably, but we prefer the term knowledge when implying a long-term accumulation process or accumulated information in our memory system (i.e., knowledge is learned information or information that has just been integrated into existing knowledge).

4. Some theories posit that awareness of a knowledge gap is triggered only by external stimuli (Berlyne, 1960) but this is not always the case. Beyond perception of the external environment, humans are also equipped with the capacity to immerse themselves with internal thoughts (Singer, 1975), being able to find or even create uncertainty in their internal knowledge structure.

5. Note that the extended processes discussed here have been implicitly or partially implemented by some of the existing quantitative models in a variety of forms. For example, the idea that knowledge acquisition influences the reward value for new exploration is implicit in some models (e.g., Stojic et al., 2019). The importance of incorporating perceived skills as an element of the model has been recognized in the literature on reinforcement learning (‘competence-based models’; Mirolli & Baldassarre, 2013). The aim of the proposed framework is to make this point explicit and upfront as the critical issue to connect different research traditions.

6. We suggest that the same issue applies to the concept of “intrinsic motivation.” Several researchers have argued how interest (or curiosity) is distinct from intrinsic motivation (Deci, 1992; Hidi, 2000; Renninger, 2000), but we do not see a clear consensus in the literature. From our perspective, intrinsic motivation is also a concept that is psychologically constructed to explain people’s behavior based on internal rewarding feeling (Murayama et al., 2019), and therefore, like curiosity and interest, it does not exist in the reward-learning processes (i.e., Figure 2 does not have any elements of “intrinsic motivation”). Because of the constructed nature of the concept, people can have different definitions and no-one can objectively prove which is correct. In other words, there is an inherent difficulty in reaching an agreed-upon definition.

7. Perceived competence (or perceived skills) is determined in relation to task demand. That is, the same individual can have different perceived competence depending on the difficulty of the task.

References

de Abril, I. M., & Kanai, R. (2018). A unified strategy for implementing curiosity and empowerment driven reinforcement learning. *ArXiv Preprint ArXiv:1806.06505*.

Afifi, W. A., & Weiner, J. L. (2006). Seeking information about sexual health: Applying the theory of motivated information management. Human *Communication Research*, *32*, 35–57. doi: 10.1111/j.1468-2958.2006.00002.x

Alexander, P. A., & Jetton, T. L. (1996). The role of importance and interest in the processing of text. *Educational Psychology Review, 8*(1), 89-121. doi:10.1007/BF01761832

Allport, G. W. (1961). *Pattern and growth in personality.* (pp. xiv, 593). Holt, Reinhart & Winston.

Amabile, T. M., Hennessey, B. A., & Grossman, B. S. (1986). Social influences on creativity: The effects of contracted-for reward. *Journal of Personality and Social Psychology, 50*(1), 14-23. doi:10.1037/0022-3514.50.1.14

Andersen, M. M., Schjoedt, U., Price, H., Rosas, F. E., Scrivner, C., & Clasen, M. (2020). Playing With Fear: A Field Study in Recreational Horror. *Psychological Science*, *31(12)*, 1497–1510. https://doi.org/10.1177/0956797620972116

Aslan, S., Fastrich, G., Donnellan, E., Jones, D., & Murayama, K. (2022). People's naïve belief about curiosity and interest: A qualitative study. *Under review*

Austin, J. T., & Vancouver, J. B. (1996). Goal constructs in psychology: Structure, process, and content. *Psychological Bulletin*, *120*, 338–375.

Back, M. D., Baumert, A., Denissen, J. J. A., Hartung, F.-M., Penke, L., Schmukle, S. C., . . . Wrzus, C. (2011). PERSOC: A unified framework for understanding the dynamic interplay of personality and social relationships. *European Journal of Personality, 25*(2), 90-107. doi:10.1002/per.811

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Educational Psychology Review, 84*, 191-215.

Bandura, A., Ross, D., & Ross, S. A. (1963). Imitation of film-mediated aggressive models. *The Journal of Abnormal and Social Psychology, 66*(1), 3-11. doi:10.1037/h0048687

Baranes, A., & Oudeyer, P.-Y. (2013). Active learning of inverse models with intrinsically motivated goal exploration in robots. *Robotics and Autonomous Systems, 61*(1), 49-73. doi:10.1016/j.robot.2012.05.008

Baranes, A., Oudeyer, P.-Y., & Gottlieb, J. (2014). The effects of task difficulty, novelty and the size of the search space on intrinsically motivated exploration. *Frontiers in Neuroscience, 8*. doi:10.3389/fnins.2014.00317

Barto, A. G., Singh, S., & Chentanez, N. (2004). Intrinsically motivated learning of hierarchical collections of skills. *Proceedings of the 3rd International Conference on Development and Learning*, 112–19.

Bennett, D., Bode, S., Brydevall, M., Warren, H., & Murawski, C. (2016). Intrinsic valuation of information in decision making under uncertainty. *Plos Computational Biology, 12*(7), e1005020. doi:10.1371/journal.pcbi.1005020

Berlyne, D. E. (1954). A theory of human curiosity. *British Journal of Psychology, 45*, 180-191.

Berlyne, D. E. (1960). *Conflict, arousal, and curiosity.* New York: McGraw-Hill.

Berlyne, D. E. (1971). *Aesthetics and psychobiology.* New York: Appleton-Century-Crofts.

Bernstein, M. R. (1955). Relationship between Interest and Reading Comprehension. *The Journal of Educational Research*, *49*(4), 283–288. https://doi.org/10.1080/00220671.1955.10882283

Berridge, K. C. (2000). Reward learning: Reinforcement, incentives, and expectations. In *Psychology of Learning and Motivation* (Vol. 40, pp. 223-278): Academic Press.

Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research Reviews, 28*(3), 309-369. doi:https://doi.org/10.1016/S0165-0173(98)00019-8

Bevins, R. A. (2001). Novelty seeking and reward: Implications for the study of high-risk behaviors. *Current Directions in Psychological Science*, *10(6)*, 189-193. doi:10.1111/1467-8721.00146

Blanchard, Tommy C., Hayden, Benjamin Y., & Bromberg-Martin, Ethan S. (2014). Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. *Neuron*, *85(3)*, 602-614. doi:10.1016/j.neuron.2014.12.050

Bong, M. (2001). Between- and within-domain relations of academic motivation among middle and high school students: Self-efficacy, task-value, and achievement goals. *Journal of Educational Psychology*, *93*, 23-34.

Braver, T. S., Krug, M. K., Chiew, K. S., Kool, W., Clement, N. J., Adcock, A., . . . Somerville, L. H. (2014). Mechanisms of motivation-cognition interaction: Challenges and opportunities. *Cognitive Affective & Behavioral Neuroscience, 14 (3),* 443-472. Doi: 10.3758/s13415-014-0300-0

Brief, A. P., & Aldag, R. J. (1977). The intrinsic-extrinsic dichotomy: Toward conceptual clarity. *Academy of Management Review, 2*(3), 496-500. doi:10.5465/amr.1977.4281861

Bromberg-Martin, E. S., & Hikosaka, O. (2009). Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron, 63*(1), 119-126. doi:https://doi.org/10.1016/j.neuron.2009.06.009

Brydevall, M., Bennett, D., Murawski, C., & Bode, S. (2018). The neural encoding of information prediction errors during non-instrumental information seeking. *Scientific Reports*, *8(1)*, 6134. doi:10.1038/s41598-018-24566-x

Buehler, R., & McFarland, C. (2001). Intensity bias in affective forecasting: The role of temporal focus. *Personality and Social Psychology Bulletin, 27*(11), 1480-1493. doi:10.1177/01461672012711009

Burgess, L. G., Riddell, P. M., Fancourt, A., & Murayama, K. (2018). The Influence of social contagion within education: A motivational perspective. *Mind, Brain, and Education, 12*(4), 164-174. doi:doi:10.1111/mbe.12178

Burke, C. J., Tobler, P. N., Baddeley, M., & Schultz, W. (2010). Neural mechanisms of observational learning. *Proceedings of the National Academy of Sciences of the United States of America, 107*(32), 14431-14436. doi:10.1073/pnas.1003111107

Campbell, D. P., & Borgen, F. H. (1999). Holland's Theory and the Development of Interest Inventories. *Journal of Vocational Behavior, 55*(1), 86-101. doi:https://doi.org/10.1006/jvbe.1999.1699

Caspi, A., & Roberts, B. W. (2001). Personality Development Across the Life Course:The Argument for Change and Continuity. *Psychological Inquiry, 12*(2), 49-66. doi:10.1207/S15327965PLI1202_01

Cerasoli, C. P., Nicklin, J. M., & Ford, M. T. (2014). Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis. *Psychological Bulletin, 140*(4), 980-1008. doi:10.1037/a0035661

Charpentier, C. J., Bromberg-Martin, E. S., & Sharot, T. (2018). Valuation of knowledge and ignorance in mesolimbic reward circuitry. *Proceedings of the National Academy of Sciences, 115*(31), E7255-E7264. doi:10.1073/pnas.1800547115

Christianson, N. H., Sizemore Blevins, A., & Bassett, D. S. (2020). Architecture and evolution of semantic networks in mathematics texts. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *476(2239)*, 20190741. https://doi.org/10.1098/rspa.2019.0741

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, *107(2)*, 261.

Csikszentmihalyi, M. (1975). *Beyond Boredom and Anxiety.* Jossey-Bass Publishers.

Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience.* New York, NY: Harper and Row.

Daniel, R., & Pollmann, S. (2010). Comparing the Neural Basis of Monetary Reward and Cognitive Feedback during Information-Integration Category Learning. *Journal of Neuroscience, 30*(1), 47-55. doi:10.1523/jneurosci.2205-09.2010

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*, 876–879.

Day, H. I. (1971). *The measurement of specific curiosity.* New York: Hold, Rinehart, & Winston.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology, 18*, 105-115.

Deci, E. L. (1992). The relation of interest to the motivation of behavior: A self-determination theory perspective. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development (p. 43–70)*. Lawrence Erlbaum Associates, Inc.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*, 627-668.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York: Plenum.

Delgado, M. R. (2007). Reward-Related Responses in the Human Striatum. *Annals of the New York Academy of Sciences, 1104*(1), 70-88. doi:doi:10.1196/annals.1390.002

Dickinson, A., & Balleine, B. (2002). The role of learning in the operation of motivational systems. In H. Pashler & R. Gallistel (Eds.), *Stevens' handbook of experimental psychology, learning, motivation and emotion (vol. 3).* (pp. 497-534).

Di Domenico, S. I., & Ryan, R. M. (2017). The Emerging Neuroscience of Intrinsic Motivation: A New Frontier in Self-Determination Research. *Frontiers in Human Neuroscience*, *11*. https://doi.org/10.3389/fnhum.2017.00145

Donnellan, E., Sakaki, M., & Murayama, K. (2021). From curiosity to interest: Accumulated knowledge supports long-term persistence of information-seeking behaviour. In I. Cogliati Dezza, E. Schulz & C. M. Wu (Eds.) (forthcoming) *The drive for knowledge: the science of human information-seeking*. Cambridge University Press.

Donnellan, E., Sumeyye, Fastrich, G. M., & Murayama, K. (2020). *How are curiosity and interest different? Naïve Bayes classification of people's naïve belief*. https://doi.org/10.31234/osf.io/697gk

Dubey, R., & Griffiths, T. L. (2020). Reconciling novelty and complexity through a rational analysis of curiosity. *Psychological Review*, *127*(3), 455–476. https://doi.org/10.1037/rev0000175

Dubey, R., Mehta, H., & Lombrozo, T. (2021). Curiosity Is Contagious: A Social Influence Intervention to Induce Curiosity. *Cognitive Science*, *45(2)*, e12937. https://doi.org/10.1111/cogs.12937

Duzel, E., Bunzeck, N., Guitart-Masip, M., & Duzel, S. (2010). NOvelty-related Motivation of Anticipation and exploration by Dopamine (NOMAD): Implications for healthy aging. *Neuroscience and Biobehavioral Reviews, 34*(5), 660-669. doi:10.1016/j.neubiorev.2009.08.006

Dweck, C. S. (1986). Motivational process affects learning. *American Psychologist*, *41*, 1010–1018.

Dyer, L., & Parker, D. F. (1975). Classifying outcomes in work motivation research: An examination of the intrinsic-extrinsic dichotomy. *Journal of Applied Psychology, 60*(4), 455-458. doi:10.1037/h0076914

Eisenberger, R., & Cameron, J. (1996). Detrimental effects of reward: Reality or mith? *American Psychologist*, *51*, 1153–1166.

Eisenberger, R., Pierce, W. D., & Cameron, J. (1999). Effects of reward on intrinsic motivation-Negative, neutral, and positive: Comment on Deci, Koestner, and Ryan(1999). *Psychological Bulletin, 125*, 677-691.

Eisenberger, R., & Selbst, M. (1994). Does reward increase or decrease creativity? *Journal of Personality and Social Psychology, 66*(6), 1116-1127. doi:10.1037/0022-3514.66.6.1116

Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist*, *34*, 169–189.

Elliot, A. J., & Harackiewicz, J. M. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology*, *70*, 461–475.

Ellsworth, P. C. (2003). Confusion, concentration, and other emotions of interest: Commentary on Rozin and Cohen (2003). *Emotion, 3*(1), 81-85. doi:10.1037/1528-3542.3.1.81

Engel, S. (2011). Children's Need to Know: Curiosity in Schools. *Harvard Educational Review, 81*(4), 625-645,784.

Fastrich, G. M., Kerr, T., Castel, A. D., & Murayama, K. (2018). The role of interest in memory for trivia questions: An investigation with a large-scale database. *Motivation Science*, *4(3)*, 227-250. doi:10.1037/mot0000087

Fastrich, G. M., & Murayama, K. (2020). Development of interest and role of choice during sequential knowledge Acquisition. *AERA Open*, *6*(2), 2332858420929981. https://doi.org/10.1177/2332858420929981

Fayn, K., Silvia, P. J., Dejonckheere, E., Verdonck, S., & Kuppens, P. (2019). Confused or curious? Openness/intellect predicts more positive interest-confusion relations. *Journal of Personality and Social Psychology*, *117*, 1016-1033. doi:10.1037/pspp0000257

Fayn, K., Silvia, P. J., MacCann, C., & Tiliopoulos, N. (2017). Interested in Different Things or in Different Ways? *Journal of Individual Differences, 38*(4), 265-273. doi:10.1027/1614-0001/a000243

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *The Journal of Abnormal and Social Psychology, 58*(2), 203-210. doi:10.1037/h0041593

FitzGibbon, L., Komiya, A., & Murayama, K. (2020). The lure of counterfactual curiosity: People incur a cost to experience regret. *Psychological Science*

Fleeson, W. (2001). Toward a structure- and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology*, *80*(6), 1011–1027. https://doi.org/10.1037/0022-3514.80.6.1011

Frenzel, A. C., Pekrun, R., Dicke, A.-L., & Goetz, T. (2012). Beyond Quantitative Decline: Conceptual Shifts in Adolescents' Development of Interest in Mathematics. *Developmental Psychology, 48*, 1069-1082.

Frey, B. S., & Jegen, R. (2001). Motivation crowding theory. *Journal of Economic Surveys, 15*, 589-611.

Gagné, M., & Deci, E. L. (2005). Self-determination theory and work motivation. *Journal of Organizational Behavior*, *26*(4), 331–362. https://doi.org/10.1002/job.322

Gigerenzer, G., & Garcia-Retamero, R. (2017). Cassandra's regret: The psychology of not wanting to know. *Psychological Review, 124*(2), 179-196. doi:10.1037/rev0000055

Goetz, T., Frenzel, A. C., Pekrun, R., Hall, N. C., & Lüdtke, O. (2007). Between-and within-domain relations of students' academic emotions. *Journal of Educational Psychology, 99*(4), 715.

Golman, R., Gurney, N., & Loewenstein, G. (2020). Information gaps for risk and ambiguity. *Psychological Review*, No Pagination Specified-No Pagination Specified. https://doi.org/10.1037/rev0000252

Gopnik, A., Griffiths, T. L., & Lucas, C. G. (2015). When Younger Learners Can Be Better (or at Least More Open-Minded) Than Older Ones. *Current Directions in Psychological Science, 24*(2), 87-92. doi:10.1177/0963721414556653

Gottfried, A. E., Fleming, J. S., & Gottfried, A. W. (2001). Continuity of academic intrinsic motivation from childhood through late adolescence: A longitudinal study. *Journal of Educational Psychology, 93*(1), 3-13. doi:10.1037/0022-0663.93.1.3

Gottfried, A. E., Marcoulides, G. A., Gottfried, A. W., & Oliver, P. H. (2013). Longitudinal Pathways From Math Intrinsic Motivation and Achievement to Math Course Accomplishments and Educational Attainment. *Journal of Research on Educational Effectiveness, 6*(1), 68-92. doi:10.1080/19345747.2012.698376

Gottlieb, J., Lopes, M., & Oudeyer, P.-Y. (2016). Motivated cognition: Neural and computational mechanisms of curiosity, attention and intrinsic motivation. In S.-i. Kim, J. Reeve, & M. Bong (Eds.), *Advances in Motivation and Achievement (Vol. 19 Recent Developments in Neuroscience Research on Human Motivation)* (pp. 149-172). Bingley, UK: Emerald Publishing Limited.

Gottlieb, J., & Oudeyer, P.-Y. (2018). Towards a neuroscience of active sampling and curiosity. *Nature Reviews Neuroscience, 19*(12), 758-770. doi:10.1038/s41583-018-0078-0

Grant, S., Kajii, A., & Polak, B. (1998). Intrinsic Preference for Information. *Journal of Economic Theory, 83*(2), 233-259. doi:https://doi.org/10.1006/jeth.1996.2458

Grossnickle, E. M. (2016). Disentangling Curiosity: Dimensionality, Definitions, and Distinctions from Interest in Educational Contexts. *Educational Psychology Review*, *28*(1), 23–60. https://doi.org/10.1007/s10648-014-9294-y

Gruber, M. J., Gelman, Bernard D., & Ranganath, C. (2014). States of Curiosity Modulate Hippocampus-Dependent Learning via the Dopaminergic Circuit. *Neuron, 84*(2), 486-496. doi:http://dx.doi.org/10.1016/j.neuron.2014.08.060

Gruber, M. J., & Ranganath, C. (2019). How Curiosity Enhances Hippocampus-Dependent Memory: The Prediction, Appraisal, Curiosity, and Exploration (PACE) Framework. *Trends in Cognitive Sciences*, *23*(12), 1014–1025. https://doi.org/10.1016/j.tics.2019.10.003

Harackiewicz, J. M., Manderlink, G., & Sansone, C. (1984). Rewarding pinball wizardry: Effects of evaluation and cue value on intrinsic interest. *Journal of Personality and Social Psychology, 47*(2), 287-300.

Hatano, A., Ogulmus, C., Shigemasu, H., & Murayama, K. (2020). *Thinking about thinking: People Underestimate Intrinsically Motivating Experiences of Waiting* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/r6mde

Heath, C. (1999). On the social psychology of agency relationships: Lay theories of motivation overemphasize extrinsic incentives. *Organizational Behavior and Human Decision Processes, 78*(1), 25-62.

Hertwig, R., & Engel, C. (2016). Homo Ignorans:Deliberately Choosing Not to Know. *Perspectives on Psychological Science, 11*(3), 359-372. doi:10.1177/1745691616635594

Hewett, R., & Conway, N. (2016). The undermining effect revisited: The salience of everyday verbal rewards and self-determined motivation. *Journal of Organizational Behavior, 37*(3), 436-455. doi:doi:10.1002/job.2051

Hidi, S. (2000). An interest researcher's perspective: The effects of extrinsic and intrinsic factors on motivation. In C. Sansone & J. M. Harackiewicz (Eds.), Intrinsic and extrinsic motivation: The search for optimal motivation and performance (p. 309–339). Academic Press. https://doi.org/10.1016/B978-012619070-0/50033-7

Hidi, S. (2016). Revisiting the Role of Rewards in Motivation and Learning: Implications of Neuroscientific Research. *Educational Psychology Review, 28*(1), 61-93. doi:10.1007/s10648-015-9307-5

Hidi, S., & Harackiewicz, J. M. (2000). Motivating the academically unmotivated: A critical issue for the 21st century. *Review of Educational Research, 70*, 151-179.

Hidi, S., & Renninger, K. A. (2006). The Four-Phase Model of Interest Development. *Educational Psychologist, 41*(2), 111-127. doi:http://dx.doi.org/10.1207/s15326985ep4102_4

Hidi, S. E., & Renninger, K. A. (2019). Interest development and its relation to curiosity: Needed neuroscientific research. *Educational Psychology Review*. doi:10.1007/s10648-019-09491-3

Hoff, K. A., Briley, D. A., Wee, C. J. M., & Rounds, J. (2018). Normative changes in interests from adolescence to adulthood: A meta-analysis of longitudinal studies. *Psychological Bulletin, 144*(4), 426-451. doi:10.1037/bul0000140

Holland, J. L. (1985). *Making vocational choices: A theory of vocational personalities and work environments.* Prentice Hall.

Howard-Jones, P. A., & Jay, T. (2016). Reward, learning and games. *Current Opinion in Behavioral Sciences*, *10*, 65-72. doi:https://doi.org/10.1016/j.cobeha.2016.04.015

Hsee, C. K., & Ruan, B. (2016). The Pandora Effect: The Power and Peril of Curiosity. *Psychological Science*. doi:10.1177/0956797616631733

Hughes, R. N. (2007). Neotic preferences in laboratory rodents: Issues, assessment and substrates. *Neuroscience & Biobehavioral Reviews, 31*(3), 441-464. doi:https://doi.org/10.1016/j.neubiorev.2006.11.004

Hull, C. L. (1943). *Principles of behavior: an introduction to behavior theory*: Oxford, England: Appleton-Century.

Hull, C. L. (1952). *A behavior system; an introduction to behavior theory concerning the individual organism*. New Haven, CT, US: Yale University Press.

Hulleman, C. S., Thoman, D. B., Dicke, A.-L., & Harackiewicz, J. M. (2017). The Promotion and Development of Interest: The Importance of Perceived Values. In P. A. O'Keefe & J. M. Harackiewicz (Eds.), *The Science of Interest* (pp. 189-208). Cham: Springer International Publishing.

Izard, C. E. (1977). *Human emotions.* New York, NY: Plenum Press.

Izard, C. E. (2009). Emotion Theory and Research: Highlights, Unanswered Questions, and Emerging Issues. *Annual review of psychology, 60*, 1-25. doi:10.1146/annurev.psych.60.110707.163539

Izuma, K., Saito, D. N., & Sadato, N. (2008). Processing of social and monetary rewards in the human striatum. *Neuron*, *58*, 284-294. doi: 10.1016/j.neuron.2008.03.020.

Jach, H. K., DeYoung, C. G., & Smillie, L. D. (2021). Why do people see information? The role of personality traits and situation perception. *Journal of Experimental Psychology: General*

Jepma, M., Verdonschot, R. G., van Steenbergen, H., Rombouts, S., & Nieuwenhuis, S. (2012). Neural mechanisms underlying the induction and relief of perceptual curiosity. *Frontiers in Behavioral Neuroscience, 6*. doi:10.3389/fnbeh.2012.00005

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. doi:http://dx.doi.org/10.1016/j.dr.2012.04.002

Jones, R. M., Somerville, L. H., Li, J., Ruberry, E. J., Libby, V., Glover, G., . . . Casey, B. J. (2011). Behavioral and Neural Properties of Social Reinforcement Learning. *The Journal of Neuroscience, 31*(37), 13039-13045. doi:10.1523/jneurosci.2972-11.2011

Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T.-y., & 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. doi:10.1111/j.1467-9280.2009.02402.x

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA, US: The MIT Press.

Kashdan, T. B. (2004). Curiosity. In C. Peterson & M. E. P. Seligman (Eds.), *Character strengths and virtues: A handbook and classification* (pp. 125-141). Washington, DC: American Psychological Association and Oxford University Press.

Kashdan, T. B., Stiksma, M. C., Disabato, D. J., McKnight, P. E., Bekier, J., Kaji, J., & Lazarus, R. (2018). The five-dimensional curiosity scale: Capturing the bandwidth of curiosity and identifying four unique subgroups of curious people. *Journal of Research in Personality, 73*, 130-149. doi:https://doi.org/10.1016/j.jrp.2017.11.011

Kashdan, T. B. & Steger, M. (2007). Curiosity and pathways to well-being and meaning in life: Traits, states, and everyday behaviors. *Motivation and Emotion*, *31*(3), 159–173. https://doi.org/10.1007/s11031-007-9068-7

Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA, US: The MIT Press.

Kidd, C., & Hayden, Benjamin Y. (2015). The Psychology and Neuroscience of Curiosity. *Neuron, 88*(3), 449-460. doi:http://dx.doi.org/10.1016/j.neuron.2015.09.010

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks Effect: Human Infants Allocate Attention to Visual Sequences That Are Neither Too Simple Nor Too Complex. *PLOS ONE, 7*(5), e36399. doi:10.1371/journal.pone.0036399

Kish, G. B. (1955). Learning when the onset of illumination is used as the reinforcing stimulus. *Journal of Comparative and Physiological Psychology, 48*(4), 261-264. doi:10.1037/h0040782

Kohn, A. (1993). *Punished by rewards: The trouble with gold stars, incentive plans, A's, praise, and other bribes*. Boston, MA: Houghton, Mifflin and Company.

Krapp, A. (1999). Interest, motivation and learning: An educational-psychological perspective. *European Journal of Psychology of Education*, *14*(1), 23–40.

Krapp, A. (2000). 5 - Interest and Human Development During Adolescence: An Educational-Psychological Approach. In J. Heckhausen (Ed.), *Advances in Psychology* (Vol. 131, pp. 109–128). North-Holland. https://doi.org/10.1016/S0166-4115(00)80008-4

Kruglanski, A. W. (1975). The endogenous-exogenous partition in attribution theory. *Psychological Review, 82*(6), 387-406. doi:10.1037/0033-295X.82.6.387

Kruglanski, A. W., Fishbach, A., Woolley, K., Bélanger, J. J., Chernikova, M., Molinario, E., & Pierro, A. (2018). A structural model of intrinsic motivation: On the psychology of means-ends fusion. *Psychological Review*, *125(2)*, 165–182. https://doi.org/10.1037/rev0000095

Kuvaas, B., Buch, R., Weibel, A., Dysvik, A., & Nerstad, C. G. L. (2017). Do intrinsic and extrinsic motivation relate differently to employee outcomes? *Journal of Economic Psychology*, *61*, 244-258. doi:https://doi.org/10.1016/j.joep.2017.05.004

Langsdorf, P., Izard, C. E., Rayias, M., & Hembree, E. A. (1983). Interest expression, visual fixation, and heart rate changes in 2- and 8-month-old infants. *Developmental Psychology, 19*(3), 375-386. doi:10.1037/0012-1649.19.3.375

Lanzetta, J. T., & Driscoll, J. M. (1966). Preference for information about an uncertain but unavoidable outcome. *Journal of Personality and Social Psychology, 3*(1), 96-102. doi:10.1037/h0022674

Lau, J. K. L., Ozono, H., Kuratomi, K., Komiya, A., & Murayama, K. (2020). Shared striatal activity in decisions to satisfy curiosity and hunger at the risk of electric shocks. *Nature Human Behaviour*, *4(5)*, 531–543. https://doi.org/10.1038/s41562-020-0848-3

Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining childrens' intrinsic interest with extrinsic reward: Test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology, 28*, 129-137.

Ligneul, R., Mermillod, M., & Morisseau, T. (2018). From relief to surprise: Dual control of epistemic curiosity in the human brain. *NeuroImage, 181*, 490-500. doi:https://doi.org/10.1016/j.neuroimage.2018.07.038

Linder, D. E., Cooper, J., & Jones, E. E. (1967). Decision freedom as a determinant of the role of incentive magnitude in attitude change. *Journal of Personality and Social Psychology, 6*(3), 245-254. doi:10.1037/h0021220

Litman, J. A. (2005). Curiosity and the pleasures of learning: Wanting and liking new information. *Cognition and Emotion, 19*(6), 793-814. doi:10.1080/02699930541000101

Litman, J. A. (2008). Interest and deprivation factors of epistemic curiosity. *Personality and Individual Differences, 44*(7), 1585-1595. doi:http://dx.doi.org/10.1016/j.paid.2008.01.014

Litman, J.A. (2019). Curiosity: Nature, dimensionality, and determinants, In K.A. Renninger & S.E. Hidi (Eds.) *The Cambridge Handbook of Motivation and Learning* (pp. 418-442). New York, NY: Cambridge University Press.

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

Litman J. A., & Jimerson T. L. (2004) The measurement of curiosity as a feeling of deprivation. *Journal of Personality Assessment*, *82*, 147-157.

Litman, J. A. & Lunsford, G. D. (2010), Incurious motives to seek information about potential threats. *European Journal of Personality*, *24*, 639-655. doi:10.1002/per.766Litman, J. A., & Pezzo, M. V. (2007). Dimensionality of interpersonal curiosity. *Personality and Individual Differences, 43*(6), 1448-1459. doi:https://doi.org/10.1016/j.paid.2007.04.021

Litman, J. A., & Spielberger, C. D. (2003). Measuring Epistemic Curiosity and Its Diversive and Specific Components. *Journal of Personality Assessment, 80*(1), 75-86. doi:10.1207/S15327752JPA8001_16

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

Lydon-Staley, D. M., Zhou, D., Blevins, A. S., Zurn, P., & Bassett, D. S. (2020). Hunters, busybodies and the knowledge network building associated with deprivation curiosity. *Nature Human Behaviour*, 1–10. https://doi.org/10.1038/s41562-020-00985-7

Lydon - Staley, D. M., Zurn, P., & Bassett, D. S. (2020). Within-person variability in curiosity during daily life and associations with well-being. *Journal of Personality*, *88*(4), 625–641. https://doi.org/10.1111/jopy.12515

Ma, Q., Jin, J., Meng, L., & Shen, Q. (2014). The dark side of monetary incentive: how does extrinsic reward crowd out intrinsic motivation. *NeuroReport, 25*(3), 194-198. doi:10.1097/wnr.0000000000000113

MacGregor, D. (1960). *The human side of enterprise* (Vol. 21): New York.

Markey, A., & Loewenstein, G. (2014). Curiosity. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International Handbook of Emotions in Education* (pp. 246-264). New York, NY: Routledge.

Marsh, H. W., & Shavelson, R. J. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist, 20*, 107-123.

Marvin, C. B., & Shohamy, D. (2016). Curiosity and reward: Valence predicts choice and information prediction errors enhance learning. *Journal of Experimental Psychology: General, 145*(3), 266-272. doi:10.1037/xge0000140

Matyjek, M., Meliss, S., Dziobek, I., & Murayama, K. (2020). A Multidimensional View on Social and Non-Social Rewards. *Frontiers in Psychiatry*, *11*. https://doi.org/10.3389/fpsyt.2020.00818

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60*(2), 175–215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

McGillivray, S., Murayama, K., & Castel, A. D. (2015). Thirst for knowledge: The effects of curiosity and interest on memory in younger and older adults. *Psychology and Aging, 30*(4), 835-841.

Metcalfe, J., Schwartz, B. L., & Bloom, P. A. (2017). The tip-of-the-tongue state and curiosity. *Cognitive Research: Principles and Implications*, *2*, 31-31. doi:10.1186/s41235-017-0065-4

Miele, D. B., & Scholer, A. A. (2018). The Role of Metamotivational Monitoring in Motivation Regulation. *Educational Psychologist, 53*(1), 1-21. doi:10.1080/00461520.2017.1371601

Miller, D. T., & Ratner, R. K. (1998). The disparity between the actual and assumed power of self-interest. *Journal of Personality and Social Psychology, 74*(1), 53-62. doi:10.1037/0022-3514.74.1.53

Mirolli, M., & Baldassarre, G. (2013). Functions and mechanisms of intrinsic motivations. In G. Baldassarre & M. Mirolli (Eds.), *Intrinsically motivated*

learninig in natural and artificial systems. (pp. 49-72). New York, NY: Springer.

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review, 102*(2), 246-268. doi:10.1037/0033-295x.102.2.246

Mobbs, D., Yu, R. J., Meyer, M., Passamonti, L., Seymour, B., Calder, A. J., . . . Dalgleish, T. (2009). A Key Role for Similarity in Vicarious Reward. *Science, 324*(5929), 900-900. doi:10.1126/science.1170539

Moneta, G. B., & Csikszentmihalyi, M. (1996). The Effect of Perceived Challenges and Skills on the Quality of Subjective Experience. *Journal of Personality*, *64(2)*, 275–310. https://doi.org/10.1111/j.1467-6494.1996.tb00512.x

Mortillaro, M., Mehu, M., & Scherer, K. R. (2011). Subtly Different Positive Emotions Can Be Distinguished by Their Facial Expressions. *Social Psychological and Personality Science, 2*(3), 262-271. doi:10.1177/1948550610389080

Mount, M. K., Barrick, M. R., Scullen, S. M., & Rounds, J. (2005). Higher-order dimensions of the Big Five personality traits and the Big Six vocational interest types.. *Personnel Psychology, 58*(2), 447-478. doi:doi:10.1111/j.1744-6570.2005.00468.x

Muis, K. R., Chevrier, M., & Singh, C. A. (2018). The Role of Epistemic Emotions in Personal Epistemology and Self-Regulated Learning. *Educational Psychologist, 53*(3), 165-184. doi:10.1080/00461520.2017.1421465

Murayama, K. (2014). Knowing your motivation: Metamotivation. *Annual Review of Japanese Child Psychology*, 112-116.

Murayama, K. (2019). Neuroscientific and Psychological Approaches to Incentives: Commonality and Multifaceted Views. In K. A. Renninger & S. E. Hidi (Eds.), *The Cambridge Handbook of Motivation and Learning* (pp. 141–162). Cambridge University Press.

Murayama, K. (2019, June 2). A reward-learning framework of autonomous knowledge acquisition: An integrated account of curiosity, interest, and intrinsic-extrinsic rewards. https://doi.org/10.31219/osf.io/zey4k

Murayama, K. (2021, October 29). A reward-learning framework of knowledge acquisition: An integrated account of curiosity, interest, and intrinsic-extrinsic rewards. Retrieved from osf.io/rsxuv

Murayama, K., & Elliot, A. J. (2011). Achievement motivation and memory: Achievement goals differentially influence immediate and delayed remember–know recognition memory. *Personality and Social Psychology Bulletin, 37*(10), 1339-1348. doi:10.1177/0146167211410575

Murayama, K., FitzGibbon, L., & Sakaki, M. (2019). Process Account of Curiosity and Interest: A Reward-Learning Perspective. *Educational Psychology Review*, *31(4)*, 875–895. https://doi.org/10.1007/s10648-019-09499-9

Murayama, K., Kitagami, S., Tanaka, A., & Raw, J. A. L. (2017). People's Naiveté About How Extrinsic Rewards Influence Intrinsic Motivation. *Motivation Science*, No Pagination Specified. doi:10.1037/mot0000040

Murayama, K., Kuratomi, K., Johnsen, L., Kitagami, S., & Hatano, A. (2018). People Underestimate their Capability to Motivate Themselves without Performance-based Extrinsic Incentives.

Murayama, K., Matsumoto, M., Izuma, K., & Matsumoto, K. (2010). Neural basis of the undermining effect of monetary reward on intrinsic motivation. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 107*(49), 20911-20916.

Murayama, K., Matsumoto, M., Izuma, K., Sugiura, A., Ryan, R. M., Deci, E. L., & Matsumoto, K. (2015). How Self-Determined Choice Facilitates Performance: A Key Role of the Ventromedial Prefrontal Cortex. *Cerebral Cortex, 25*(5), 1241-1251. doi:10.1093/cercor/bht317

Murayama, K., Pekrun, R., Lichtenfeld, S., & vom Hofe, R. (2013). Predicting Long-Term Growth in Students' Mathematics Achievement: The Unique

Contributions of Motivation and Cognitive Strategies. *Child Development, 84*(4), 1475-1490. doi:10.1111/cdev.12036

Mussel, P. (2013). Intellect: A Theoretical Framework for Personality Traits Related to Intellectual Achievements. *Journal of Personality and Social Psychology, 104*(5), 885-906. doi:10.1037/a0031918

Nielsen, L., Knutson, B., & Carstensen, L. L. (2008). Affect dynamics, affective forecasting, and aging. *Emotion, 8*(3), 318-330. doi:10.1037/1528-3542.8.3.318

O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: Insights from neuroimaging. *Current Opinion in Neurobiology, 14*, 769-776.

O'Keefe, P., & Linnenbrink-Garcia, L. (2014). *The Role of Interest in Optimizing Performance and Self-Regulation* (Vol. 58).

Oosterwijk, S. (2017). Choosing the negative: A behavioral demonstration of morbid curiosity. *PLOS ONE, 12*(7), e0178399. doi:10.1371/journal.pone.0178399

Oosterwijk, S., Snoek, L., Tekoppele, J., Engelbert, L., & Scholte, H. S. (2019). Choosing to view morbid information involves reward circuitry. bioRxiv, 795120. doi:10.1101/795120

Oudeyer, P.-Y., & Kaplan, F. (2009). What is intrinsic motivation? A typology of computational approaches. *Frontiers in Neurorobotics, 1*. doi:10.3389/neuro.12.006.2007

Ozono, H., Komiya, A., Kuratomi, K., Hatano, A., Fastrich, G., Raw, J. A. L., Haffey, A., Meliss, S., Lau, J. K. L., & Murayama, K. (2020). Magic Curiosity Arousing Tricks (MagicCATs): A novel stimulus collection to induce epistemic emotions. *Behavior Research Methods*. https://doi.org/10.3758/s13428-020-01431-2

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions.* New York: Oxford University Press.

Pathak, D., Agrawal, P., Efros, A. A., & Darrell, T. (2018). Curiosity-driven exploration by self-supervised prediction. *arXiv preprint arXiv:1705.05363, 2017.*

Piaget, J. (1952). *The origins of intelligence in children*. New York, NY, US: W W Norton & Co.

Pintrich, P. R., & DeGroot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*, 33-40.

Prenzel, M. (1992). Selective persistence of interest. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development.* (pp. 71-98). Hillsdale, NJ: Erlbaum.

Reed, P., Mitchell, C., & Nokes, T. (1996). Intrinsic reinforcing properties of putatively neutral stimuli in an instrumental two-lever discrimination task. *Animal Learning & Behavior, 24*(1), 38-45. doi:10.3758/bf03198952

Reeve, J., & Lee, W. (2019). A neuroscientific perspective on basic psychological needs. *Journal of Personality*, *87(1)*, 102–114. https://doi.org/10.1111/jopy.12390

Renninger, K. A. (2000). Chapter 13 - Individual interest and its implications for understanding intrinsic motivation. In C. Sansone & J. M. Harackiewicz (Eds.), *Intrinsic and Extrinsic Motivation* (pp. 373-404). San Diego: Academic Press.

Renninger, K. A., & Hidi, S. (2002). Student interest and achievement: Developmental issues raised by a case study. In A. Wigfield & J. Eccles (Eds.), *Development of achievement motivation: A volume in the educational psychology series.* (pp. 173-195). San Diego, CA: Academic.

Renninger, K. A., & Hidi, S. (2016). *The power of interest for motivation and engagement.* New York, NY: Routledge.

Robinson, O. C., Demetre, J. D., & Litman, J. A. (2017). Adult life stage and crisis as predictors of curiosity and authenticity:Testing inferences from Erikson's

lifespan theory. *International Journal of Behavioral Development, 41*(3), 426-431. doi:10.1177/0165025416645201

Rodriguez Cabrero, J. A. M., Zhu, J.-Q., & Ludvig, E. A. (2019). Costly curiosity: People pay a price to resolve an uncertain gamble early. *Behavioural Processes, 160*, 20-25. doi:https://doi.org/10.1016/j.beproc.2018.12.015

Ross, M. (1975). Salience of reward and intrinsic motivation. *Journal of Personality and Social Psychology, 32*(2), 245-254. doi:10.1037/0022-3514.32.2.245

Rotgans, J. I., & Schmidt, H. G. (2011). The role of teachers in facilitating situational interest in an active-learning classroom. *Teaching and Teacher Education*, *27*(1), 37–42. https://doi.org/10.1016/j.tate.2010.06.025

Ruan, B., Hsee, C. K., & Lu, Z. Y. (2018). The Teasing Effect: An Underappreciated Benefit of Creating and Resolving an Uncertainty. *Journal of Marketing Research, 55*(4), 556-570. doi:10.1509/jmr.15.0346

Rushworth, M. F. S., Mars, R. B., & Summerfield, C. (2009). General mechanisms for making decisions? *Current Opinion in Neurobiology, 19*(1), 75-83.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110(1)*, 145–172. https://doi.org/10.1037/0033-295X.110.1.145

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*, 54-67.

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. New York, NY, US: Guilford Press.

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, *61*, 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Sakaki, M., Yagi, A., & Murayama, K. (2018). Curiosity in old age: A possible key to achieving adaptive aging. *Neuroscience & Biobehavioral Reviews, 88*, 106-116. doi:https://doi.org/10.1016/j.neubiorev.2018.03.007

Sansone, C. (1989). Competence feedback, task feedback, and intrinsic interest: An examination of process and context. *Journal of Experimental Social Psychology*, *25(4)*, 343–361. https://doi.org/10.1016/0022-1031(89)90027-9

Sansone, C., & Thoman, D. B. (2005). Interest as the Missing Motivator in Self-Regulation. *European Psychologist, 10*(3), 175-186. doi:10.1027/1016-9040.10.3.175

Sansone, C., Thoman, D. B., & Smith, J. L. (2010). Interest and Self-Regulation: Understanding individual variability in choices, efforts, and persistence over time. In R. Hoyle (Ed.), *Handbook of Personality and Self-Regulation* (pp. 192-217). Malden, MA: Blackwell.

Sansone, C., Weir, C., Harpster, L., & Morgan, C. (1992). Once a boring task always a boring task? Interest as a self-regulatory mechanism. *Journal of Personality and Social Psychology, 63*(3), 379-390. doi:10.1037/0022-3514.63.3.379

Schembri, M., Mirolli, M., & Baldassarre, G. (2007). Evolving internal reinforcers for an intrinsically motivated reinforcement-learning robot. *2007 IEEE 6th International Conference on Development and Learning*, 282–287.

Schiefele, U. (1991). Interest, Learning, and Motivation. *Educational Psychologist*, *26*(3–4), 299–323. https://doi.org/10.1080/00461520.1991.9653136

Schmidhuber, J. (1991). *A possibility for implementing curiosity and boredom in model-building neural controllers*. Paper presented at the Proceedings of the first international conference on simulation of adaptive behavior on From animals to animats, Paris, France.

Schmitt, F. F., & Lahroodi, R. (2008). The epistemic value of curiosity. *Educational Theory, 58*(2), 125-148.

Scholer, A. A., Miele, D. B., Murayama, K., & Fujita, K. (2018). New Directions in Self-Regulation: The Role of Metamotivational Beliefs. *Current Directions in Psychological Science, 27*(6), 437-442. doi:10.1177/0963721418790549

Schultz, E., Wu, M. C., Ruggeri, A., & Bjorn, M. (2019). Searching for rewards like a child means less generalization and more directed exploration. *BioRxiv (*https://doi.org/10.1101/473975*)*.

Shah, J. Y., & Kruglanski, A. W. (2008). Structural dynamics: The challenge of change in goal systems. In J. Y. Shah & W. L. Gardner (Eds.), *Handbook of motivation science*. (pp. 217–229). Guilford Press.

Sharot, T., & Sunstein, C. R. (2020). How people decide what they want to know. *Nature Human Behaviour*, *4*(1), 14–19. https://doi.org/10.1038/s41562-019-0793-1

Sheldon, K. M., Gunz, A., Nichols, C. P., & Ferguson, Y. (2010). Extrinsic Value Orientation and Affective Forecasting: Overestimating the Rewards, Underestimating the Costs. *Journal of Personality, 78*(1), 149-178. doi:doi:10.1111/j.1467-6494.2009.00612.x

Shelhamer, E., Mahmoudieh, P., Argus, M., & Darrell, T. (2017). Loss is its own Reward: Self-Supervision for Reinforcement Learning. *ArXiv:1612.07307 [Cs]*. http://arxiv.org/abs/1612.07307

Shin, D. D., & Kim, S.-I. (2019). Homo Curious: Curious or interested? *Educational Psychology Review*. doi:10.1007/s10648-019-09497-x

Sizemore, A. E., Karuza, E. A., Giusti, C., & Bassett, D. S. (2018). Knowledge gaps in the early growth of semantic feature networks. *Nature Human Behaviour*, *2(9)*, 682–692. https://doi.org/10.1038/s41562-018-0422-4

Silvia, P. J. (2001). Interest and interests: The psychology of constructive capriciousness. *Review of general psychology, 5*(3), 270-290. doi:10.1037/1089-2680.5.3.270

Silvia, P. J. (2005). What Is Interesting? Exploring the Appraisal Structure of Interest. *Emotion, 5*(1), 89-102. doi:10.1037/1528-3542.5.1.89

Silvia, P. J. (2006). *Exploring the psychology of interest*: Oxford University Press, New York, NY.

Silvia, P. J. (2008). Interest--The curious emotion. *Current Directions in Psychological Science, 17*(1), 57-60. doi:http://dx.doi.org/10.1111/j.1467-8721.2008.00548.x

Silvia, P. J., & Christensen, A. P. (2020). Looking up at the curious personality: Individual differences in curiosity and openness to experience. *Current Opinion in Behavioral Sciences*, *35*, 1–6. https://doi.org/10.1016/j.cobeha.2020.05.013

Silvia, P. J., Henson, R. A., & Templin, J. L. (2009). Are the sources of interest the same for everyone? Using multilevel mixture models to explore individual differences in appraisal structures. *Cognition and Emotion, 23*(7), 1389-1406. doi:http://dx.doi.org/10.1080/02699930902850528

Singer, J. L. (1975). *Daydreaming and fantasy*. London: Routledge.

Smith, C. A., & Ellsworth, P. C. (1988). Shades of Joy: Patterns of Appraisal Differentiating Pleasant Emotions AU - Ellsworth, Phoebe C. *Cognition and Emotion, 2*(4), 301-331. doi:10.1080/02699938808412702

Somerville, L. H., Sasse, S. F., Garrad, M. C., Drysdale, A. T., Abi Akar, N., Insel, C., & Wilson, R. C. (2017). Charting the expansion of strategic exploratory behavior during adolescence. *Journal of Experimental Psychology: General, 146*(2), 155-164. doi:10.1037/xge0000250

Spinath, B., & Steinmayr, R. (2008). Longitudinal analysis of intrinsic motivation and competence beliefs: Is there a relation over time? *Child Development, 79*(5), 1555-1569.

Steyvers, M., & Tenenbaum, J. B. (2005). The Large-Scale Structure of Semantic Networks: Statistical Analyses and a Model of Semantic Growth. *Cognitive Science*, *29(1)*, 41–78. https://doi.org/10.1207/s15516709cog2901_3

Stojic, H., Schulz, E., Analytis, P. P., & Speekenbrink, M. (2019). It's new, but is it good? How generalization and uncertainty guide the exploration of novel options.

Su, R. (2018). The three faces of interests: An integrative review of interest research in vocational, organizational, and educational psychology. *Journal of Vocational Behavior*. doi:https://doi.org/10.1016/j.jvb.2018.10.016

Suzuki, S., Harasawa, N., Ueno, K., Gardner, J. L., Ichinohe, N., Haruno, M., . . . Nakahara, H. (2012). Learning to Simulate Others' Decisions. *Neuron, 74*(6), 1125-1137. doi:10.1016/j.neuron.2012.04.030

Tachimori, Y., Iwanaga, H., & Tahara, T. (2013). The networks from medical knowledge and clinical practice have small-world, scale-free, and hierarchical features. *Physica A: Statistical Mechanics and Its Applications*, *392*(23), 6084–6089. https://doi.org/10.1016/j.physa.2013.07.047

Ten, A., Kaushik, P., Oudeyer, P., & Gottlieb, J. (2021). Humans monitor learning progress in curiosity-driven exploration. *Nature Communications*.

Lydon-Staley, D. M., Zhou, D., Blevins, A. S., Zurn, P., & Bassett, D. S. (2020). Hunters, busybodies and the knowledge network building associated with deprivation curiosity. Nature Human Behaviour, 1–10. https://doi.org/10.1038/s41562-020-00985-7

Tang, S. H., & Hall, V. C. (1995). The overjustification effect: A meta-analysis. *Applied Cognitive Psychology, 9*(5), 365-404. doi:10.1002/acp.2350090502

Thoman, D. B., Sansone, C., & Geerling, D. (2017). The Dynamic Nature of Interest: Embedding Interest Within Self-Regulation. In P. A. O'Keefe & J. M. Harackiewicz (Eds.), *The Science of Interest* (pp. 27–47). Springer International Publishing. https://doi.org/10.1007/978-3-319-55509-6_2

Thoman, D. B., Sansone, C., Robinson, J. A., & Helm, J. L. (2020). Implicit theories of interest regulation. *Motivation Science*, *6(4)*, 321–334. https://doi.org/10.1037/mot0000160

Tobias, S. (1995). Interest and metacognitive word knowledge. *Journal of Educational Psychology, 87*(3), 399-405. doi:10.1037/0022-0663.87.3.399

Tomkins, S. S. (1962). *Affect, imagery, consciousness: Vol. I. The positive affects*. Springer.

Tracey, T. J., & Rounds, J. (1993). Evaluating Holland's and Gati's vocational-interest models: A structural meta-analysis. *Psychological Bulletin*, *113*, 229–246.

Tricomi, E., Delgado, M. R., McCandliss, B. D., McClelland, J. L., & Fiez, J. A. (2006). Performance feedback drives caudate activation in a phonological learning task. *Journal of Cognitive Neuroscience, 18*(6), 1029-1043. doi:10.1162/jocn.2006.18.6.1029

Turner Jr, S. A., & Silvia, P. J. (2006). Must interesting things be pleasant? A test of competing appraisal structures. *Emotion*, *6*(4), 670.

Twomey, K. E., & Westermann, G. (2018). Curiosity-based learning in infants: a neurocomputational approach. *Developmental Science, 21*(4), e12629. doi:10.1111/desc.12629

van Lieshout, L. L. F., Vandenbroucke, A. R. E., Müller, N. C. J., Cools, R., & de Lange, F. P. (2018). Induction and relief of curiosity elicit parietal and frontal activity. *The Journal of Neuroscience*. doi:10.1523/jneurosci.2816-17.2018

van Lieshout, L., Traast, I., de Lange, F., & Cools, R. (2019, September 24). Curiosity or savouring? Information seeking is modulated by both uncertainty and valence. https://doi.org/10.31234/osf.io/5y6pz

Vansteenkiste, M., Simons, J., Lens, W., Soenens, B., & Matos, L. (2005). Examining the motivational impact of intrinsic versus extrinsic goal framing and autonomy-supportive versus internally controlling communication style on early adolescents' academic achievement. *Child Development, 76*(2), 483-501.

von Stumm, S., Hell, B., & Chamorro-Premuzic, T. (2011). The Hungry Mind: Intellectual Curiosity Is the Third Pillar of Academic Performance.

Perspectives on Psychological Science, 6(6), 574-588. doi:10.1177/1745691611421204

Wade, S., & Kidd, C. (2019). The role of prior knowledge and curiosity in learning. *Psychonomic Bulletin & Review*, *26*, 1377-1387. doi:10.3758/s13423-019-01598-6

Walker, E. L. (1981). The Quest for the Inverted U. In H. I. Day (Ed.), *Advances in Intrinsic Motivation and Aesthetics* (pp. 39-70). Boston, MA: Springer US.

Wang, M. Z., & Hayden, B. Y. (2019). Monkeys are curious about counterfactual outcomes. *Cognition, 189*, 1-10. doi:https://doi.org/10.1016/j.cognition.2019.03.009

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, *393*(6684), 440–442. https://doi.org/10.1038/30918

Westgate, E. C., & Wilson, T. D. (2018). Boring thoughts and bored minds: The MAC model of boredom and cognitive engagement. *Psychological Review*, *125*, 689-713. doi:10.1037/rev0000097

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review, 66*(5), 297-333.

Wiersma, U. J. (1992). The effects of extrinsic rewards on intrinsic motivation: A meta-analysis.. *Journal of Occupational and Organizational Psychology, 65*, 101-114.

Wigfield, A., & Eccles, J. S. (2002). The development of competence beliefs, expectancies for success, and achievement values from childhood through adolescence. In G. Phye (Ed.), *Development of achievement motivation* (pp. 91-120). San Diego: Academic Press.

Wigfield, A., Eccles, J. S., Schiefele, U., Roeser, R. W., Davis-Kean, P., Eisenberg, N., . . . Lerner, R. M. (2006). Development of Achievement Motivation. In *Handbook of child psychology: Vol. 3, Social, emotional, and personality development (6th ed.).* (pp. 933-1002). Hoboken, NJ US: John Wiley & Sons Inc.

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans Use Directed and Random Exploration to Solve the Explore–Exploit Dilemma. *Journal of Experimental Psychology. General*, *143(6)*, 2074–2081. https://doi.org/10.1037/a0038199

Wilson, T. D., Centerbar, D. B., Kermer, D. A., & Gilbert, D. T. (2005). The Pleasures of Uncertainty: Prolonging Positive Moods in Ways People Do Not Anticipate. *Journal of Personality and Social Psychology, 88*(1), 5-21. doi:10.1037/0022-3514.88.1.5

Wilson, T. D., Wheatley, T., Meyers, J. M., Gilbert, D. T., & Axsom, D. (2000). Focalism: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology, 78*(5), 821-836. doi:10.1037/0022-3514.78.5.821

Witherby, A. E., & Carpenter, S. K. (2021). The rich-get-richer effect: Prior knowledge predicts new learning of domain-relevant information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, https://doi.org/10.1037/xlm0000996

Wittmann, B. C., Daw, N. D., Seymour, B., & Dolan, R. J. (2008). Striatal activity underlies novelty-based choice in humans. *Neuron, 58*, 967-973.

Wolters, C. A. (1998). Self-regulated learning and college students' regulation of motivation. *Journal of Educational Psychology, 90*(2), 224-235. doi:10.1037//0022-0663.90.2.224

Woolley, K., & Fishbach, A. (2015). The experience matters more than you think: People value intrinsic incentives more inside than outside an activity. *Journal of Personality and Social Psychology, 109*(6), 968-982.

Woolley, K., & Fishbach, A. (2016). For the Fun of It: Harnessing Immediate Rewards to Increase Persistence in Long-Term Goals. *Journal of Consumer Research, 42*(6), 952-966. doi:10.1093/jcr/ucv098

Wrzesniewski, A., Schwartz, B., Cong, X., Kane, M., Omar, A., & Kolditz, T. (2014). Multiple types of motives don't multiply the motivation of West Point cadets.

Proceedings of the National Academy of Sciences, 111(30), 10990-10995. doi:10.1073/pnas.1405298111

Wrzus, C., & Roberts, B. W. (2017). Processes of Personality Development in Adulthood: The TESSERA Framework. *Personality and Social Psychology Review, 21*(3), 253-277. doi:10.1177/1088868316652279

Wyckoff Jr, L. B. (1952). The role of observing responses in discrimination learning. Part I. *Psychological Review, 59*(6), 431-442. doi:10.1037/h0053932

Zentall, T. R. (2005). A Within-trial Contrast Effect and its Implications for Several Social Psychological Phenomena. *International Journal of Comparative Psychology, 18*(4), 273-297.

Ziegler, M., Cengia, A., Mussel, P., & Gerstorf, D. (2015). Openness as a buffer against cognitive decline: The Openness-Fluid-Crystallized-Intelligence (OFCI) model applied to late adulthood. *Psychology and Aging, 30*(3), 573-588. doi:10.1037/a0039493

Zuckerman, M. (1979). *Sensation seeking: Beyond the optimallevel of arousal*. Hillsdale: Lawrence Erlbaum.

Zurn, P., & Bassett, D. S. (2018). On curiosity: A fundamental aspect of personality, a practice of network growth. *Personality Neuroscience*, *1*, e13. doi:10.1017/pen.2018.3