Word count:

- Abstracts: 119 + 170 words
- Main text: 8612 words
- References: 4173 words
- Entire Text: 13074 words

A critique of motivation constructs to explain higher-order behavior: We should unpack the black box

Kou Murayama[1,2] & Hayley Jach[1]

University of Tübingen

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

[2]Kochi University of Technology, 185 Miyanokuchi, Tosayamada, Kami-shi, Kochi, 782-8502, Japan

Other information

Kou Murayama:

k.murayama@uni-tuebingen.de; Homepage: https://motivationsciencelab.com/

Hayley Jach

hayleyjach@gmail.com

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Short Abstract

The current article critically evaluates motivation constructs explaining a wide range of higher-order behavior which have burgeoned in the literature of psychology. We argue that, while such high-level motivation constructs seemingly explain higher-order behavior quite well, they do not specify what they actually are and how they produce behavior, casting doubt on their theoretical status. To address this “black box problem”, we demonstrate the importance of specifying mental computational processes underlying such higher order motivated behavior, and argue that high-level motivation constructs actually reflect a psychological construction of the regularities of our subjective experiences and behavior. The proposed perspective opens new avenues for future theoretical development, namely the examination of how motivated behavior is *realized* through mental computational processes.

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Long Abstract

The constructs of motivation (or needs, motives, etc.) to explain higher-order behavior have burgeoned in psychology. In this article, we critically evaluate such high-level motivation constructs that many researchers define as causal determinants of behavior. We identify a fundamental issue with this predominant view of motivation, which we called the *black-box problem.* Specifically, high-level motivation constructs have been considered as causally instigating a wide range of higher-order behavior, but this does not explain what they actually are or how behavioral tendencies are generated. The black box problem inevitably makes the construct ill-defined and jeopardizes its theoretical status. To address the problem, we discuss the importance of mental computational processes underlying motivated behavior. Critically, from this perspective, motivation is not a unitary construct that causes a wide range of higher-order behavior --- it is an emergent property that people construe through the regularities of subjective experiences and behavior. The proposed perspective opens new avenues for future theoretical development, i.e., the examination of how motivated behavior is *realized* through mental computational processes.

Key words: causality, computational modeling, essentialism, incentives, measurement model, meta-theory, intrinsic motivation, reward

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

1. Introduction

From the inception of psychology and across the ensuing century, constructs of motivation have played a critical role in explaining human behavior (Hull, 1943; McDougall, 1909). Researchers have considered motivation one of the most essential ingredients in our mind, addressing the fundamental question of why people initiate certain behaviors in the first place (Kanfer & Chen, 2016). Initially the constructs of motivation were mostly proposed for basic behavior such as eating (e.g., hunger drive) or mating (e.g., sex drive). However, later years have seen increasing use of motivation constructs to explain higher-order behavior (though such use was observed in early years; e.g., Murray, 1938). Nowadays, there is a plethora of “high-level motivation constructs” in psychology, including (but not limited to) the need for competence, relatedness, and autonomy (Deci & Ryan, 1985), the need to belong (Baumeister & Leary, 1995), self-affirmation motive (Steele, 1988), desire for status (Anderson et al., 2015), self-enhancement motive (Sedikides & Strube, 1997), achievement motive (McClelland et al., 1976), and intrinsic-extrinsic motivation (Deci & Ryan, 1985).

In the current opinion paper[1], we provide a critical analysis of these motivation constructs to explain higher-order human behavior. Specifically, we cast doubt on the theoretical status of high-level motivation in the sense of a construct that directly influences complex behavior. Rather, we contend that such high-level motivation is a subjective construal or emergent property of underlying mental computational processes which determine behavior. From this “psychological construction” perspective, we clarify both strengths and weaknesses of high-level motivation constructs, and offer a new avenue of research which has attracted almost no attention in the past: theoretical analysis of how motivation is *realized* through mental computational processes.

2. Motivation for higher-order behavior: Definitions and clarifications

2-1. Definition

[1] The preliminary idea of the manuscript was already discussed in a short essay (Murayama, 2022b).

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

The definition of motivation varies across different times and fields (e.g., Kleinginna & Kleinginna, 1981; Madsen, 1974), but one common definition is that motivation *energizes* and *directs* behavior (Lewin, 1942; Locke & Latham, 2004; Reeve, 2017; Simpson & Balsam, 2016; VandenBos & American Psychological Association, 2007; Weiner, 1992). Energization means that motivation instigates or initiates action or behavior (Elliot, 2023), which shall be called the “spring to action” of behavior (James, 1890), a force that impels behavior (Descartes, 1955), or “energy behind our actions” (Wigfield et al., 2021). On the other hand, direction means that it guides and channel behavior in a certain way. The former aspect of motivation is often referred to as “motives” (Anderson et al., 2015; Atkinson & Raynor, 1978; McClelland et al., 1976) or “needs” (Deci & Ryan, 1985; Dweck, 2017; Maslow, 1943; Stevens & Fiske, 1995). The latter aspect of motivation is referred to as “goals” or “values” (Austin & Vancouver, 1996; Eccles & Wigfield, 2002; Elliot & Fryer, 2008; Fishbach & Ferguson, 2007). Combined together, motivation is conceptualized as a determinant of a certain set of behaviors (Figure 1A).

*** Insert Figure 1 around here ***

Our critical analysis of motivation constructs mainly concerns the former aspect, the function of energization (see also Hinde, 1960, for another critique). Regardless of one’s theory of motivation, motivation is almost always used to explain the initiation (or the intention of the initiation) of behavior; as such, energization is often regarded as the definitive aspect of motivation (e.g., Madsen, 1974). The direction aspect of motivation is also often considered a fundamental aspect of motivation, but we view this aspect as somewhat subsumed within the energization aspect (see also Elliot, 2023). In fact, most of the existent constructs of high-level motives or needs emphasize energization in terms of what they are motivated *for* (i.e., direction). For example, an achievement motive represents the motivation toward a high standard of excellence (McClelland et al., 1976) and a need for autonomy directs people toward fulfilling a sense of agency (Ryan & Deci, 2017). It is difficult to imagine motives or needs that instigate behavior in a completely nonspecific manner, other than a few limited examples (e.g., general

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Pavlovian instrumental transfer; Corbit & Balleine, 2005). We will briefly revisit the direction aspect of motivation in a later section.

The concept of motivation to explain basic human behavior (e.g., mating, consumption of foods) was initially subject to criticisms concerning its operationalization (e.g., Koch, 1941). Although these points overlap with our criticism to some degree, our main criticism is aimed at the motivation constructs that explain a broad range of higher-order behavior, which we shall call high-level motivation constructs.

2-2. An example

To render our criticism concrete and easy to understand, we first show one example of how high-level motivation constructs are used to explain behavior: the *need for competence* (Ryan & Deci, 2020). The need for competence is based on the concept of competence proposed by White (1959). According to White (1959), competence reflects the organism's capacity to effectively interact with its environment. Extending this idea, Deci and Ryan (1985) proposed that humans have a basic psychological need for competence, which is defined as people's motivation to experience feelings of mastery and success. We use the need for competence as an example simply because it is one of the most accepted high-level motivation constructs in the field (19,800 hits in Google Scholar in September, 2023), and recent theoretical progress has made it possible to instantiate our point (Murayama, 2022).

Humans and animals often exhibit behavior that seems to be aimed at mastering the environment. Research showed that monkeys are engaged with solving puzzles (Harlow, 1949). Humans have the capacity to engage in learning for a prolonged period of time (Hidi & Renninber, 2019). Humans and animals also have a tendency to explore the environment without clear rewards (Berlyne, 1966) and a tendency to seek positive feedback (Elliot & Moller, 2003). None of these behaviors can be explained by basic motivation such as a hunger drive, and are considered to be caused by the need for competence (Deci & Ryan, 1985). Using this construct, we can explain the behavior in the following way: "We have a tendency to explore because we have a basic motivation to master the environment," and, "People can sustain their commitment to an activity

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

because people have fundamental motivation to seek mastery." The satisfaction of this need for competence is theorized to enhance one's intrinsic motivation (i.e., motivation to work on a task without relying on explicit incentives), and to lead to many positive long-term outcomes such as higher well-being (Ryan & Deci, 2017, 2020).

3. Fundamental challenges for high-level motivation constructs

3.1 The Black-box Problem

As the definition of motivation as an energizer and director of behavior attests, researchers typically regard motivation as the initial cause (i.e., origin) of a certain set of behaviors (Figure 1A). Of course, researchers have assumed that motivation constructs are influenced by many external factors, such as environmental changes, learning, socialization, and development, and that they have genetic origins as well as neural bases (McClelland, 1987). In early days, motivation was conceptualized as an "intervening variable" (Hull, 1943), meaning that it has both external antecedents (e.g., deprivation of food) and outcomes (e.g., increased response). However, many motivation constructs are regarded as the origin of behavior in the sense that they are the internal variable which is supposed to generate the willpower to initiate the action in the first place.

This property of motivation constructs is particularly useful for researchers to understand higher-order behavior. As noted in the example above, we can explain exploratory behavior by proposing that we have a basic motivation to master the environment. Crucially, however, the high-level motivation construct does not truly explain *what* it is or *how* this behavioral tendency is generated (for historic arguments, see Bindra, 1959; Koch, 1956). Instead, motivation is like a black box, where the process that generates the behavior is unknown. By supposing motivation to explain behavior, we implicitly fall prey to the so-called "motivational homunculus" problem (see also Gladwin et al., 2011). That is, to explain how motivated behavior is generated, we posit a construct that works as a "generator" of motivated behavior, but this logic may suffer from the issue of infinite regress (Kenny, 1971). The real danger here is that,

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

because the constructs seemingly explain the set of higher-order behaviors well, we may think that all questions are answered, and soon stop investigating what these constructs are and how they work.

Historically speaking, there was a time when researchers tried to eliminate the black-box property of motivation constructs by proposing physiological or biological causes. For example, Hull's concept of drive such as hunger driver or thirst drive was directly linked to the physiological deficits of food or water (Hull, 1943). There are also contemporary theories that connect some motivation constructs with simple physiological or biological factors (e.g., testosterone and power motivation; Schultheiss et al., 1999). However, recent studies suggest that such simple one-to-one correspondence between a physiological factor and motivation is not plausible to explain motivation constructs for higher-order complex behavior (e.g., Kim, 2013; Murayama et al., 2017; Steinman et al., 2019). We believe that the high-level motivation constructs partly gained popularity because they paralleled motivation for basic behavior. With the analogy of motivation for food, the statement, "We have a tendency to do *X* because we have a fundamental motivation for it," sounds intuitive and convincing. However, when applied to higher-order behavior, motivation constructs are much more likely to suffer from the black-box problem.

Note that our argument differs from the issue of circular explanation that is often discussed in the classical literature (e.g., Bindra, 1959; Seward, 1939; i.e., motivation constructs explain a particular type of behavior by arguing that people have a motivation for that behavior). Motivation constructs do have great utility in that they can make generalizable predictions (Berridge, 2004) --- for example, by supposing humans have a need for competence, one can predict that humans perform a range of epistemic behaviors, even if these behaviors were not part of the original observation. The constructs also make our explanation parsimonious --- now we can conceptualize these behaviors as all manifestations of the single motivational construct of the need to belong. By assuming general properties for motivation constructs, we can even make novel predictions for behavior (Baumeister & Leary, 1995). Our point is rather that

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

motivation is considered a useful explanatory variable (i.e., *explanans*), but it does not have explanatory variables itself (i.e., *explanandum*).

3.2. Lack of Consensus on the Definition of High-level Motivation Constructs

The problem of the motivational black box gives rise to another critical issue: Challenges in defining high-level motivation constructs. Because motivation constructs are created to explain a certain set of behaviors without specifying their internal properties, they can be defined only in terms of the behaviors explained by the constructs (see Figure 1A). As a result, there is always room for ambiguity when one tries to define them based on their internal properties. In other words, high-level motivation constructs have an inherent challenge for precise definition due to their black-box property.

This issue has manifested in various forms in the literature of motivation. In early years, researchers tried to create a comprehensive list of human needs (McDougall, 1926; Murray, 1938), but there was always a question of how we could be certain that two similar needs were distinct and not the same, or whether certain needs were or were not fundamental (Pittman & Zeigler, 2007). In recent years, this issue has often been discussed in the context of jingle-jangle fallacies of motivation constructs (e.g., Bong, 1996; Pekrun, 2023; Pekrun & Marsh, 2022), where the same construct label is used to denote different constructs or different construct labels are used for the same construct (Kelley, 1927). Taking the example of need for competence again, the construct is considered to be a source of intrinsic motivation (Deci & Ryan, 1985, 2020; i.e. if this need is satisfied, intrinsic motivation increases). Does this mean they are separate constructs? Or is need for competence a constituent part of intrinsic motivation? There are also other constructs related to need for competence. For example, self-efficacy is a belief in one's capacity to competently control the environment (Bandura, 1997) and is clearly related to need for competence. Perceived control (Skinner, 1996), self-esteem (Baumeister, 1993) and self-concept (Marsh & Shavelson, 1985) are also similar constructs clearly connected to need for competence. Are they different constructs and, if not, what are the relationships? Great effort has been devoted to resolving these issues, and the work helped researchers to deepen our thoughts about these motivation

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

constructs. However, given the black-box property, it is virtually impossible to make a conclusive judgement on the difference between constructs (see also Murayama et al., 2019).

This issue of defining high-level motivation constructs is especially problematic when researchers want to make a causal inference. For example, in survey studies, researchers use a variety of designs (e.g., longitudinal study) to estimate a causal effect of a motivation construct on outcome variables. But does that mean that motivation has a causal effect? To make a causal inference, we estimate the effect when motivation is (hypothetically) “intervened on”, holding other factors constant. However, given the black-box property of high-level motivation constructs, researchers often have difficulty in judging whether or not some potential controlling variables are the inherent property of the motivation constructs (see Eronen, 2020; Hernán & VanderWeele, 2011). For example, think of a situation in which researchers want to learn the causal effect of (satisfaction of) need for competence on math exam performance using longitudinal survey data. One may want to treat self-esteem as a controlling variable, but one could also argue that self-esteem is a constituent part of need for competence. In such cases, controlling for self-esteem does not make sense. But such a decision is often difficult because the constructs are underidentified. Generally speaking, when the target construct is not unambiguously defined, we can never make a solid causal inference from empirical data (Rohrer & Murayama, 2021).

Some may argue that the issue could be addressed empirically. For example, researchers often use the strategy of testing “incremental validity”, in which motivation construct A has predictive power over an outcome variable above and beyond a similar motivation construct B (Smith et al., 2003). Some other researchers use factor analysis and show that items representing motivation A and B form distinctive factors (Byrne, 2001; based on certain statistical criteria). Positive results from these analyses often lead researchers to conclude that the two motivation constructs “are overlapping but distinct.” However, the analysis does not directly answer the question of definitions, because the evidence simply shows that the two *measurements* are assessing

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

something different: it is mute to exactly how the two *theoretical constructs* are differently defined. In fact, after such a conclusion, it is often not entirely clear what it really means that two motivational constructs are overlapping but still distinct.

4. Specifying Mental Computational Processes underlying Motivation: A Potential Solution

4-1. High-level Motivation as a Psychological Construction

We offer an alternative perspective on high-level motivation constructs. Specifically, we propose that we should not view high-level motivation constructs as the original causal determinant of behavior. Rather, we argue that such high-level motivation constructs are an emergent property of underlying mental computational processes (Figure 1B). To make sense of these emergent properties, humans construe the construct of motivation. In other words, high-level motivation constructs are a consequence of psychological construction.

We define mental computational processes as concrete internal mechanisms which produce behavior (see also Marr, 1982). Consider a robot who behaves like humans. The robot employs particular computational mechanisms to process external input (i.e., sensory input) and decide on actions (i.e., output). Oftentimes some stored information in the robot (memory) plays a critical role for this computation. We call all of these mechanisms mental computational processes. Of course, humans are not robots. While these mental computational processes determine people's behavior, humans also have subjective experiences such as positive and negative feelings, and these subjective experiences should be influenced by mental computational processes (LeDoux, 2014)[2]. Figure 1B provides a schematic picture.

[2] As indicated by the figure, subjective experiences can exert impact on mental computational processes. There has been a long discussion on whether this is true or not (e.g., Sheldon, 2022; Wegner, 2004) but our argument holds regardless of the standpoint on the matter. The key point is that these effects are, if any, mediated by mental computational processes.

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Importantly, humans are capable of recognizing regularities in and creating mental categories from their own behaviors and subjective experiences (Reeder, 2009). Motivation may be a convenient term to explain these categories. Many studies suggest that we are indeed naturally inclined to infer motivations or intentions from a variety of observations (e.g., Baillargeon et al., 2015). As a result, if we have a tendency to be affiliated with certain social groups, for example, people may be naturally convinced that we have an affiliative or social motivation. However, such inference does not necessarily mean that social motivation is itself represented in our mental computational processes --- instead, social motivation could be a consequence of interpreting and categorizing the regularities that exist in behavioral patterns and subjective experiences. As such, motivation is the subjective interpretation of or label for an emergent property arising from underlying mental computational processes.

From this alternative perspective, the key solution to the black-box problem is simple: To unpack the black box. Specifically, researchers on motivation should take further steps to unravel the mental computational processes underlying high-level motivation constructs. In the following section, we demonstrate how this principle can be applied to the motivation construct need for competence.

4.2 A case for reward learning models of information-seeking

We propose that the black box of need for competence can begin to be unpacked with reward-learning models of information seeking behavior (e.g., FitzGibbon et al., 2020; Gottlieb et al., 2013; Gruber & Ranganath, 2019; Marvin & Shohamy, 2016). In our view, information seeking is an excellent lens from which to unpack some types of motivation, including need for competence, because the act of seeking information describes effort taken to acquire knowledge, a fundamental process to master the environment. We are not claiming that the reward-learning model is the best model to instantiate need for competence with mental computational processes; there are several alternatives (e.g., Patankar et al., 2022) and there are many different versions of reward-learning models (e.g., Oudeyer & Kaplan, 2009). However, our purpose is not to compare these models. We use this model simply for demonstration purposes as it

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

provides a useful example of how our perspective can explain certain types of motivation constructs.

Although there are many different versions of reward-learning models to explain human information-seeking behavior, a common assumption is that information is an intrinsic reward. Murayama (2022c; see also Murayama et al., 2019) summarized and expanded on these common aspects of reward-learning models in the “reward-learning framework of knowledge acquisition” (Figure 2). According to the framework, when an agent identifies some uncertainty in its knowledge (often called “knowledge gap”), the agent computes the expected rewarding value of upcoming information, and if it is deemed valuable, the agent initiates information-seeking behavior. When the agent successfully acquires the information, the agent experiences a positive rewarding feeling, which in turn strengthens the value of the same sort of information. This is a well-known reinforcement principle (Hull, 1943) but the critical point is that the framework assumes that information itself can have rewarding value. There are different algorithms proposed to accurately quantify the rewarding value of information (e.g., uncertainty; Bennett et al., 2016; Lieshout et al., 2018). Furthermore, Murayama (2022c) argued that acquired information is consolidated into the existing knowledge base, and this expanded knowledge could prompt the agent to become more aware of further knowledge gaps (i.e., “the more we know about a topic, the more likely we realize that there are things that we do not know”). As a result, this system creates a positive feedback loop, sustaining long-lasting information-seeking behavior.

*** Insert Figure 2 around here ***

The framework aims to provide a rough summary of current information-seeking models. In these models, we can see several mental computational processes included, such as assessment of knowledge gap, computation of expected reward value of information, selection of information, integration of the information into existing knowledge, and so on. These computations do not normally operate consciously, but

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

rather implicitly (Murayama et al., 2019). Some of these information-seeking models are conceptual (e.g., Gruber & Ranganath, 2019; Jach et al., 2022). Some other researchers propose computational models to accurately describe people's information-seeking behavior (for a review, see Baldassarre & Mirolli, 2013) and some other models are even implemented in robots with the aim to replicate people's (especially infants' or young children's) information-seeking behavior (e.g., Baranes & Oudeyer, 2013). This means that, by implementing these mental computational processes, we can create an agent that actively seeks information to expand its knowledge.

Critically, the agent which implements the reward-learning models actively and continuously searches for information that it does not know to expand its knowledge, *as if* it has the motivations for mastery and competence, despite there being no need for competence featured in the mental computational processes (i.e., there are no boxes or parameters directly representing need for competence in the model). Need for competence thus appears as an emergent property of reward-learning models of information-seeking behavior. Several other motivational concepts can be explained in a similar way. For example, the motivational concept of interest often refers to people's enduring tendency to engage in particular learning content over time (Hidi & Renninger, 2019; Renninger & Hidi, 2016). The agent can realize this enduring information-seeking behavior by incorporating their knowledge base and resultant positive feedback loop into the system (Murayama, 2022c; Murayama et al., 2019). The agent also appears "intrinsically motivated" in the sense that it actively searches its environment without extrinsic incentives.

Of course, there are many motivational constructs that cannot be explained by the presented reward-learning models (e.g., need to belong). Additionally, while reward-learning models are dominant in the fields of cognitive science and neuroscience, this is not the only way to specify mental computational processing of motivated behavior. But the example is intended to demonstrate that high-level motivation construct can be a consequence of the subjective construction from behavioral regularities.

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

We used need for competence and reward-learning models as an illustrative example, but there are other potential examples which demonstrate the point that higher-level motivation constructs can be an emergent property of mental computational processes. For example, Shultz and Lepper (1996) proposed a connectionist model to explain various experimental findings of cognitive dissonance (e.g., Festinger & Carlsmith, 1959). These findings are often explained by positing that humans have a fundamental motive to maintain cognitive consistency or reduce cognitive dissonance (Festinger, 1957; Heider, 1958). However, the model that Festinger and Carlsmith (1959) proposed explained the experimental findings without explicitly supposing such motivation. O'Reilly (2020) built a multi-layered connectionist model (inspired by neuroscientific findings) and argued that the model could explain motivated behavior (i.e., the dynamic nature of goal-directed behavior) without explicitly incorporating motivation into the model.

4.3 Strengths of Considering Mental Computational Processes

By specifying the mental computational processes underlying higher-order motivated behavior, high-level motivation constructs are no longer black boxes. Instead, they clearly explain *what* the motivation construct is and *how* this behavioral tendency is generated. Importantly, the proposed perspective explicitly refutes the idea that high-level motivation constructs *themselves* cause wide-ranging higher-order human behavior (see Figure 1A). Rather, our behavior is governed by the mental computational processes, which form a collective dynamic system of interacting elements. Different types of higher-order behavior (e.g., exploring the environment, sustaining epistemic engagement, seeking competence-relevant information) are the consequences of the integration or parts of this collective system, not the consequence of a unitary construct of motivation.

The proposed perspective also neatly sidesteps the fundamental challenge of defining high-level motivation constructs. This is because it is the mental computational processes, not the motivation constructs themselves, that are necessary to understand human behaviour. For example, as we presented in the previous section, a description

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

of the mental computational processes that explain how people decide to seek information over time is sufficient to explain the behaviors related to need for competence, intrinsic motivation, and interest. There is no further need to discuss which components of this process represent need for competence, intrinsic motivation, or interest, because we already explained the mechanism (see also Kidd & Hayden, 2015; Murayama et al., 2019). The proposed perspective indicates that the priority should be given to the understanding the underlying computational mechanisms, not discussing the boundaries of (inherently ill-defined) constructs.

An important benefit of the proposed perspective is that it provides a different way of theorizing about and examining motivation: Describing *how* a high-level motivation construct is *realized* by mental computational processes (for a discussion on the merit of this approach in psychology, see van Rooij & Baggio, 2021). Traditionally, researchers start by positing a specific high-level motivation construct and develop a theory by specifying the factors related to the construct (e.g., personality traits, well-being). As a consequence, empirical research has been mainly interested in examining external antecedents (e.g., family environment) and outcomes (e.g., well-being) of motivation (establishing the so-called 'nomological network'; Cronbach & Meehl, 1955). As detailed later, we see merit in this approach. However, this is just one way of understanding motivated behavior. An alternative process-focused approach can create new sets of research questions to further advance our understanding of motivation.

For example, we showed that the reward-learning framework of knowledge acquisition (Murayama, 2022c) can also provide a theory of how need for competence is realized through reward-learning mechanisms. Once such a theory is developed, researchers can empirically test its validity by examining various parts of the mental processes (e.g., "Does a knowledge gap really facilitate information-seeking behavior?"). The theory can also help us evaluate the antecedents and outcomes identified in the previous empirical literature (e.g., which component of the mental computational processes can be altered by family environment?). Furthermore, the model prompts researchers to critically analyze the assumptions underlying the model. For example, the

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

reward-learning framework assumes that information works as rewards, but the model does not specify the type of information that is perceived as rewarding. Then researchers can further theorize and empirically test the nature of information that people actively seek, further pinning down the origin of motivated behavior. In fact, this is currently a hot topic of the field (Fitzgibbon & Murayama, 2022; Gottlieb & Oudeyer, 2018; Kidd & Hayden, 2015)

4.4 Re-interpreting past empirical findings

It is important to restress that we do not disregard voluminous empirical survey and experimental studies of motivation in the past century. Much research has found that motivation, as assessed by survey questions or manipulated in experiments, is related to human behavioral and psychological outcomes. These results clearly show the usefulness of motivational constructs in predicting behavior. In fact, "motivation constructs" have been popular in psychology because they can predict various important outcomes, such as well-being, achievement scores, career choice, and so on (e.g., Robbins et al., 2004). These empirical studies can also inform researchers of potential intervention programs (e.g., Lazowski & Hulleman, 2016). However, our argument is that we should not interpret these empirical findings as evidence that high-level motivation constructs directly cause behavior.

For example, in many studies of high-level motivation constructs, researchers assess certain types of motivation using established survey questions. These survey questions often assess certain aspects of mental computational processes (e.g., the positive rewarding feeling following knowledge acquisition) or overall behavior (e.g., the frequency of active information-seeking behavior) related to the motivational construct. When the aggregated scores are related to outcomes and potential confounders are well controlled (Figure 3, top "Statistical results from data"), we tend to argue that "motivation", assessed by the survey questions, caused the outcomes (Figure 3, left, "Motivation as determinant"). This way of thinking is common, and is often strengthened by the recent proliferation of latent variable modelling (Jöreskog, 1969), which visually

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

shows path diagrams where boxes representing motivation predict or are predicted by other variables.

However, the association between aggregated scores and the outcome could happen in a different causal scenario. Specifically, these survey items which reflect parts of mental computational processes (which interact with each other) may directly influence other variables without mediating motivation constructs (Figure 3, right, “Motivation as psychological construction”). This way of viewing measurement has gained increasing attention in the methodological literature (Donnellan et al., 2022; e.g., McClure et al., 2021; VanderWeele, 2022). For example, Donnellan, Usami, and Murayama (2022) argued that such interpretations are more realistic than latent variable modelling when items represent a wide range of behavior and psychological processes, and proposed a new mixed-effects model to analyze survey data without relying on latent variables (see also Rhemtulla et al., 2020). From this perspective, aggregated survey scores capture the extent to which the mental computational process efficiently work as a whole for a particular individual or a situation (see van der Maas et al., 2006 for a similar argument in case of intelligence). Such efficiency scores are parsimonious and useful for prediction, but we do not need to assume that these scores capture the motivation construct itself.

*** Insert Figure 3 around here ***

We believe this reframing of empirical studies on motivation is particularly important because researchers often face situations in which empirical findings disagree depending on how we measure or manipulate motivation (e.g., Hulleman et al., 2010; Utman, 1997). Recent years have also seen many studies indicating that the same constructs assessed by survey questions and experimental tasks have different correlates (Dang et al., 2020; Eisenberg et al., 2019), and this has also been a recurring issue in motivation constructs (Deci et al., 1999; McClelland et al., 1989; Schultheiss et al., 2009). In such situations, we are often encouraged to accurately define the construct or establish a valid measurement. It is indeed true that we need precise definitions of

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

constructs and valid measurement, but at the same time, the argument rests on an implicit assumption that there is a single construct of motivation that determines behavior and thus the inconsistent results are problematic. From the perspective of psychological construction, on the other hand, such divergent results are natural when different assessments or manipulations tap different aspects of the mental computational processes. Having a precise process model in mind, researchers can then make more fine-grained predictions about how different types of assessments or manipulations result in different outcomes. Such a process-oriented perspective can also help researchers develop targeted interventions that aim to change specific outcomes.

5. Answering Questions

To further elaborate and avoid misunderstanding, here we answer some follow-up questions that readers might have about our perspective.

5.1. What is the distinction between basic motivation constructs and high-level motivation constructs?

The high-level motivation constructs evaluated in the current article explain a wide-range of higher-order behavior, whereas basic motivation constructs have a relatively narrow set of behaviors as explanandum (e.g., the motivation for sex to explain people's desire to copulate). However, while we made a distinction in this article for the purpose of simplicity, the distinction is continuous rather than dichotomous. We do not argue that motivation for basic behavior does not suffer from the issue of black box at all --- this is a matter of degree. Our point is that the problem is exacerbated for high-level motivation constructs.

5.2 Are mental computational processes free from motivation constructs?

No, they are not. As mental computational processes produce behaviors, we must logically assume something which initiates behavior either explicitly or implicitly. In the reward-learning framework, there is an implicit assumption that people choose to seek information that has a high reward value. Therefore, one could argue that the framework posits a "motivation to seek rewarding information", perhaps just before the box of

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

information-seeking behavior in Figure 2. Unless we understand the basic biological/physical mechanisms to produce behavior, we can never eliminate motivation-like constructs.

But our argument is not to eliminate motivation constructs from explanation. Our proposal is that theorizing mental computational processes (i.e., analyzing motivation from a different level) can provide new insights into what researchers have called “motivation” in the literature, enabling a deeper understanding of our motivated behavior. For example, both exploration behavior and long-term intellectual engagement are thought to be caused by need for competence, but there is no explanation for how these outcomes are related. By specifying the mental computational processes, we can propose that exploration behavior is a caused by the rewarding value of information, while long-term engagement is a consequence of a positive feedback loop created by expanding existing knowledge from new information (Figure 2). Both behaviors are produced by the same mental computational system (so it makes sense that they are explained by the same construct) but the way they are produced within the system is different.

We can go even further down the line to unpack motivation-like constructs in the specified mental computational processes. As noted above, the reward-learning framework implicitly assumes a motivation to seek rewarding information. But what are the basic building blocks of rewarding value for information? Is that novelty (Poli et al., 2022), entropy (Loewenstein, 1994), or Kullback-Leibler divergence (Ningombam et al., 2022)? As we do not have direct access to such information in reality, how can such metric be computed in real world? (Gottlieb, 2018) Note that as we go down the level, motivation-like constructs become narrower and narrower (or more and more specific), and at some point we may no longer feel comfortable to call it motivation. When we orient attention to an object that suddenly comes into our visual field, for example, is this a manifestation of motivation? We can in theory suppose such a motivation that causes the orientation behavior, but many researchers would say it is not necessary or useful (Murayama, 2022b).

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

In a sense our proposal is consistent with the idea that the functioning of human behavior can be understood at different levels of granularity with the lowest being the biological or physical level (e.g., Dennett, 1987; Marr, 1982; Newell, 1994; in case of motivation, see Nagengast & Trautwein, 2023). Our proposed perspective indicates that high-level motivation constructs reflect higher-level explanations whereas mental computational processes represent lower-level explanations. No level of understanding should be dismissed as “wrong” (i.e., one level of explanation should not be replaced with a lower level explanation), because they just explain the behavior for different purposes, but the problem of motivation literature is that most researchers are satisfied with higher-level explanations (i.e. supposing high-level motivation constructs to explain behavior) and little effort has been made to pursue lower-level explanations.

5.3 Has your perspective truly not been discussed in the prior motivation literature?

The idea of psychological construction is not novel at a general level --- it has been discussed in various forms in psychological and cognitive science as well as philosophy of science (Brick et al., 2022; Churchland, 1979; Dalege et al., 2016; Danziger, 1990; MacCorquodale & Meehl, 1948; Pessoa et al., 2022; Stich & Ravenscroft, 1994). Other psychological constructs such as emotions, personality, intelligence, and consciousness have also received similar scrutiny (Boag, 2018; Fiske, 2020; Lau, 2009; Russell, 2003; van der Maas et al., 2006). Recent theoretical developments in psychological network analysis (Borsboom & Cramer, 2013) have a similar root (for application to motivational constructs, see Sachisthal et al., 2019; Tamura et al., 2022). However, the construct of motivation can be uniquely positioned in this context because motivation is characterized as the causal determinant of behavior. As we showed earlier, this unique property poses fundamental challenges when theorizing or interpreting motivated behavior, and the psychological construction perspective has clear utility to circumvent the issue.

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Nevertheless, there is some (albeit limited) work related to our psychological construction perspective in the literature of motivation. For example, Bem (1967) suggested that dissonance reduction motive (Festinger, 1957) can be interpreted as the consequence of self-perception. The idea is that people are not motivated by the dissonance reduction motive, but simply act and infer attitudes by observing/interpreting what they did (i.e., attribution process). This idea is seemingly similar to the psychological construction perspective in that it does not assume a motivation to explain higher-order behavior. However, the difference is that Bem's (1967) theory uses the attribution process to explain people's actual behavior or attitude (e.g., "people changed their behavior because of attribution"). On the other hand, our proposed perspective would suggest an interplay between different mental computational processes that together produced something we would categorize as a "dissonance reduction motive". The attribution process could be part of it, but not the whole (in fact, it is not plausible that higher-order behavior is largely explained by attribution processes). Therefore, although we acknowledge that there were similar ideas in the literature of motivation, our proposed psychological construction perspective is critically distinct in that we stress the importance of specifying mental computational processes.

That said, this line of work suggests another interesting line of work for future research. Specifically, we could examine the processes of how people construct motivation from the regularities of observable behavior and subjective feelings. While early research on attribution provided many insights into this psychological construction process (e.g., Bem, 1967; Nisbett & Wilson, 1977), the proliferation of high-level motivation constructs in recent years seems to have suppressed this tradition of work. But by combining the work on mental computational processes and psychological construction processes, we may be able to achieve deeper insights into motivated behavior. For example, there is a possibility that such attributional process of motivation works as a reinforcer of motivated behavior itself, influencing mental computational processes (e.g., "I like to study because I think I have high achievement motivation"; (see Palminteri & Lebreton, 2022). Future studies could examine such interactive processes.

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

5.4 Does the criticism only apply to the energization aspect of motivation and not the direction aspect?

Our critical discussion focused mainly on the energization aspect of motivation, but the motivation is also said to additionally have the directional aspect. The directional aspect of motivation channels people's behavior in a certain way, and researchers argue that values (Eccles & Wigfield, 2002) or goals (Elliot & Fryer, 2008) play a critical role here. Does our proposed perspective have implications for the directional aspect?

As noted earlier, the distinction between energization and direction aspects in motivation is somewhat ambiguous. To clarify the distinction, Niv and colleagues (2006) introduced a normative account of motivation, by defining the directional aspect of motivation as a modulation of the mapping between reward outcome and utility (Figure 4). When one is in a state of hunger, the utility of food reward would be increased. When one is in a state of thirst, the utility of food may not be as high but the utility of water would be increased. This means that hunger motivation altered the utility of food. This is one of the commonly accepted definitions of motivation in the classic literature (Berridge, 2004; Dickinson & Balleine, 2002; Hull, 1943). By viewing motivation as the modulator of reward utility, we can independently define motivation as separate from the energization aspect (Niv et al., 2006) --- it simply changes the mapping of utility without directly causing behavior. In this respect, therefore, our criticism does not apply to the directional aspect of motivation as exemplified by hunger or thirst.

*** Insert Figure 4 around here ***

Goals and values could be considered in the same manner. Like hunger or thirst, goals and values function as the modulator of utility; when one has the value for acquiring mathematics competence, for example, they would assign high utility or experience stronger positive feelings when succeeding in a math exam than those who do not value math. When one pursues performance goals, outperforming others would give stronger positive feelings than those pursuing mastery goals. Goals and values are

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

also not viewed as initial determinants of behavior; rather, we usually think that goals and values are set in people's mind through certain mechanisms. They are often externally provided (e.g., external goal setting; Harackiewicz & Sansone, 1991; Locke & Latham, 1990; value intervention; Rozek et al., 2017) or internally generated (e.g., self-set goals; Barron & Harackiewicz, 2001; Locke, 2001; internalization of values; Renninger & Hidi, 2016). Therefore, our criticism does not directly apply to these concepts.

However, unlike hunger or thirst, theories of such higher-order constructs such as goals and values are still unclear about the mental computational mechanism of how they are internally generated (for some exceptions, see Ballard et al., 2021; Vancouver et al., 2020). For example, prominent theories, such as expectancy-value theory (Eccles & Wigfield, 2020) and the hierarchical model of achievement motivation (Elliot, 1997), identified a number of factors that influence goals or values (e.g., implicit beliefs, personal characteristics, subjective perceptions, and affective states). However, these theories do not answer *how* these factors lead to the adoption of goals and values. In fact, the hierarchical model does not identify what kind of decision-making process is involved when one adopts mastery goals over performance goals. Expectancy-value theory does not specify how value is incorporated and represented into the existing knowledge structure. Like the case of needs and motives, we believe that specifying the computational mechanisms of goal/value adoption and transformation would provide a new landscape of understanding these concepts even better. In sum, while directional aspects of motivation are indeed less immune to our criticism, higher-order motivational concepts such as goals and values still have large room which can be benefited from the proposed perspective[3].

5.5 What about the evolutionary account of motivation?

[3] One might argue that high-level motivation constructs which we discussed, such as need for competence or need to belong, can also be conceptualized as having this directional function (e.g., people add high utility for intimate social group formation). However, for the same reason described here, this does not address the essential problem of these high-level motivation constructs.

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Researchers often use evolution to justify motivation constructs, e.g., people are motivated to understand the environment because this is crucial for survival. It is true that sets of behavior explained by need for competence can have survival value. But this perspective only addresses *why* question ("why do people have motivation *X*"), not the *what* ("what is this motivation *X*?") or *how* ("how motivation *X* is realized") questions. In addition, an evolutionary perspective does not mean that we should keep motivation constructs a black box. In fact, the evolutionary perspective is open to the possibility that evolution shaped our mental computational processes in a particular manner. In other words, thinking about adaptive value and ecological constraints would help researchers specify more realistic mental computational processes (Anderson & Milson, 1989; Lieder & Griffiths, 2020). Therefore, evolution is not incompatible with our proposal that we should unpack the black box property of high-level motivation constructs --- evolution is, rather, a useful addition for our perspective.

5. Conclusion: Toward the Future of Motivation Science

In the current article, we provided a critical evaluation of motivation constructs that explain a wide range of higher-order behavior. Although such high-level motivation constructs seem to explain higher-order behavior quite well, they do not specify what they are or how they work, which we called the *black-box problem*. To address the black box problem, we highlighted the utility of specifying the mental computational processes underlying the motivated behavior. Importantly, according to this perspective, high-level motivation constructs can be understood as people's subjective construction of these mental processes, and should not be considered as the determinant of behavior. The idea of psychological construction is rather meta-theoretical; as such, it does not contradict or refute the vast number of empirical findings in the field. Nevertheless, the proposed perspective points to important avenues for future theoretical development --- theories addressing how motivation is *realized* in our mental computational processes.

We do not intend to refute the utility of existing theories of motivation. Existing theories of motivation constructs profoundly shaped our academic field, orienting

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

researchers toward important phenomena that would otherwise have been overlooked. For example, to explain human behaviors that are not driven by clear extrinsic incentives (e.g., money, food, etc.), Deci and Ryan (1985) proposed that humans have intrinsic motivation. Since the introduction of the concept of intrinsic motivation, numerous studies have examined the nature of human behavior that is not driven by extrinsic incentives, which significantly enhanced our understanding of such behavior.

At the same time, the black-box issue in the existing theories deeply constrains our theories of motivated behavior. For example, it is generally challenging to understand the relationship between different high-level motivation constructs. This is because both constructs do not specify what the construct is composed of, providing room for different ways of theorizing their relationship. Consequently, as many researchers have repeatedly indicated (Anderman, 2020; Baumeister, 2016; P. K. Murphy & Alexander, 2000), there are currently too many motivation theories and constructs, with little integration between them. Identification of mental computational processes may provide a way to understand motivated behavior more parsimoniously, because the same process can give rise to different types of motivated behavior. Notably, the reward-learning framework we presented (Figure 2) could explain the manifestation of several motivation constructs in a single framework (e.g., need for competence, intrinsic motivation, curiosity, interest; for a more comprehensive treatment, see Murayama, 2022c).

We have consistently used the need for competence as an example so that readers can follow our argument easily, and because there has been great progress in specifying computational processes underlying exploration, a behavior closely linked with need for competence (Baldassarre & Mirolli, 2013; Cogliati Dezza et al., 2022; Oudeyer, 2007). However, the idea applies to many higher-level motivation constructs. Need to belong, for example, is defined as a drive to form and maintain at least a minimum quantity of lasting, positive, and significant interpersonal relationships (Baumeister & Leary, 1995). The construct is used to explain a wide range of social behaviors, such as people's bond-forming behavior, the tendency to preserve social relations, the effective/prioritized processing of social information, social satiation (i.e.

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

reduced inclination to form new social relationship if one already has sufficient social bonds), negative emotional experiences after rejection, and so on. But the construct does not explain the processes through which these types of social behaviors unfold. For instance, the tendency to preserve social relationships may be explained by the positive feedback loop of social rewards (a positive rewarding experience due to instant social interaction) --- once a person is in a group, the person may gain constant positive social feedback from peers such that the person is not willing to leave the group. Here the term "social rewards" should still be unpacked, but it is much more narrowly defined than the need to belong. Regarding the effective processing for social stimuli, recent work using deep neural networks showed that such effective/prioritized processing (e.g., face recognition) as well as the brain functional localization of social information may emerge from our category discrimination process (Dobs et al., 2022; Kanwisher et al., 2023). This is just a simple thinking exercise (not even a hypothesis) but it opens a new avenue for theorizing or examining the need to belong.

Of course, such theorizing is not easy. This is because mental computational processes are not observable and people are typically unaware of such processes (Kihlstrom, 1987; Nisbett & Wilson, 1977). However, the field of cognitive science and other related areas have provided different ways to theorize about our mental computational processes (including reward learning models), and researchers can take advantage of these precedents as the starting point (e.g., Dayan & Abbott, 2005; Friston & Kiebel, 2009; Lieder & Griffiths, 2020). Biological constraints often inform mental computational processes of motivation (see, for example, Rolls, 2016). It is also important to add that the mental computational processes do not have to be described as quantitative formulas (i.e., computational modelling). Such a formulation is useful but neither necessary nor sufficient. The reward-learning framework we discussed above (Murayama, 2022c), for example, did not specify the computational function of how expected reward value is calculated, but it still explains how the need for competence can emerge from such a system. There are some other attempts to conceptually describe the mechanisms underlying motivated behavior (e.g., Brehm, 1999; Richter et al., 2016). In fact, the use of mathematical formulation can be challenging to describe

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

the complexity of motivated behavior, especially in real-world contexts (DeYoung & Krueger, 2020). Conversely, we can also easily trick the mathematical expression to explain motivated behavior (e.g., adding a “constant” or “bonus” in the utility value for the motivated behavior one wants to explain). In such cases, the resultant model does not really explain the origin of motivated behavior, even if the model is mathematically expressed[4]. Our point is that many roads can lead to Rome: our goal should be to describe the processes underlying motivated behavior, regardless of how this goal is achieved.

From the proposed perspective, one important avenue for future theoretical work is to understand how intra-individual states are related to stable inter-individual differences (i.e., states and traits; see also Baumeister, 2016; this is also an issue of time-scale). When motivation is defined as the determinant of behavior, motivational traits can be simply quantified and operationalized as “general strength” of motivation. Motivational states, on the other hand, can be quantified and operationalized as short within-person fluctuation of the strengths. However, mental computational processes occur at the within-person level by definition. They do not have an enduring “general strength”, and it is not obvious how stable motivational traits can be explained. To understand motivational states and traits, we need to develop a theory of mental computational processes that explicitly addresses how intra-individual processes translates into long-term development (see Dalege et al., 2016; Murayama, 2022c; see also Atkinson et al., 1977).

One big implication of the proposed perspective is that we should no longer see motivation as an inherent category. In fact, based on the reward-learning framework (Fig. 2) one can argue that intrinsically-motivated behavior is controlled by cognitive computational processes (e.g., calculation of expected reward value of new information) as well as affective experiences (i.e., rewarding feelings). Learning is also at the heart of

[4] Berridge (2023) discussed this point by comparing two different computational models of “wanting” (Smith & Read, 2022; Zhang et al., 2009).

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

this motivational processes. There has been a long tradition in psychology to distinguish several inherent categories of mental functioning such as motivation, emotion, and cognition (Danziger, 1990), and these categories formed distinct research fields with distinct theories. Using these categories, there have also been attempts to discuss how they are related. For example, some emotion theories argue that emotions have motivating functioning, subsuming motivation under the category of emotion (e.g., Arnold, 1969; Brehm, 1999; Frijda, 1986). These categories are certainly useful for academic communications, and understanding the relationship between these categories should advance our understanding of motivated behavior. However, we also suggest that these categories could constrain our thinking, potentially hindering the development of comprehensive theories to explain behavior and decision making (Murayama, 2022a). In fact, as long as we can correctly specify the mental computational processes to explain behavior, it is not that important to discuss which part of the processes are categorized as motivation, emotion, or cognition.

Halliday (1983, p. 105) states that when a psychological variable such as motivation is invented, it simply means “some phenomenon that requires explanation has been identified” (see also MacCorquodale & Meehl, 1948). We need to take this statement seriously. Motivation is not an explanation itself: it is the starting point for explanation. We hope that the proposed “unpacking the black box” perspective motivates researchers (via, of course, mental computational processes) to explore new forms of research on human motivated behavior.

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Acknowledgements

We thank Bernard Weiner and Andrew Elliot for providing comments on an earlier draft of the manuscript.

Competing interests: none.

Funding / financial support

This work was supported by the Alexander von Humboldt Foundation (Kou Murayama, the Alexander von Humboldt Professorship endowed by the German Federal Ministry of Education and Research).

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

References

Anderman, E. M. (2020). Achievement motivation theory: Balancing precision and utility. *Contemporary Educational Psychology*, *61*, 101864. https://doi.org/10.1016/j.cedpsych.2020.101864

Anderson, C., Hildreth, J. A. D., & Howland, L. (2015). Is the desire for status a fundamental human motive? A review of the empirical literature. *Psychological Bulletin*, *141*, 574–601. https://doi.org/10.1037/a0038781

Anderson, J. R., & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review*, *96*, 703–719.

Arnold, M. B. (1969). Emotion, motivation, and the limbic system. *Annals of the New York Academy of Sciences*, *159*(3), 1041–1058. https://doi.org/10.1111/j.1749-6632.1969.tb12996.x

Atkinson, J. W., Bongort, K., & Price, L. H. (1977). Explorations using computer simulation to comprehend thematic apperceptive measurement of motivation. *Motivation and Emotion*, *1*, 1–27. https://doi.org/10.1007/BF00997578

Atkinson, J. W., & Raynor, J. O. (1978). *Personality, motivation, and achievement*. Oxford, England: Hemisphere.

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

Baillargeon, R., Scott, R. M., He, Z., Sloane, S., Setoh, P., Jin, K., Wu, D., & Bian, L. (2015). Psychological and sociomoral reasoning in infancy. In *APA handbook of personality and social psychology, Volume 1: Attitudes and social cognition* (pp. 79–150). American Psychological Association. https://doi.org/10.1037/14341-003

Baldassarre, G., & Mirolli, M. (2013). *Intrinsically motivated learning in natural and arificial systems.* Springer.

Ballard, T., Palada, H., Griffin, M., & Neal, A. (2021). An Integrated Approach to Testing Dynamic, Multilevel Theory: Using Computational Models to Connect Theory, Model, and Data. *Organizational Research Methods*, *24*(2), 251–284. https://doi.org/10.1177/1094428119881209

Bandura, A. (1997). *Self-efficacy: The exercise of control.* Freeman.

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Baranes, A., & Oudeyer, P.-Y. (2013). Active learning of inverse models with intrinsically motivated goal exploration in robots. *Robot. Auton. Syst.*, *61*(1), 49–73. https://doi.org/10.1016/j.robot.2012.05.008

Barron, K. E., & Harackiewicz, J. M. (2001). Achievement goals and optimal motivation: Testing multiple goal models. *Journal of Personality and Social Psychology*, *80*, 706–722.

Baumeister, R. F. (1993). *Self-esteem: The puzzle of low self-regard*. New York, NY, US: Plenum Press.

Baumeister, R. F. (2016). Toward a general theory of motivation: Problems, challenges, opportunities, and the big picture. *Motivation and Emotion*, *40*(1), 1–10. https://doi.org/10.1007/s11031-015-9521-y

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*, 497–529.

Bem, D. J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, *74*(3), 183–200. https://doi.org/10.1037/h0024835

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. https://doi.org/10.1371/journal.pcbi.1005020

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

Berridge, K. C. (2004). Motivation concepts in behavioral neuroscience. *Physiology & Behavior*, *81*(2), 179–209. https://doi.org/10.1016/j.physbeh.2004.02.004

Bindra, D. (1959). *Motivation: A Systematic Reinterpretation* (1st Edition/1st Printing). Ronald Press Co.

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Boag, S. (2018). Personality Dynamics, Motivation, and the Logic of Explanation. *Review of General Psychology*, *22*(4), 427–436. https://doi.org/10.1037/gpr0000150

Bong, M. (1996). Problems in academic motivation research and advantages and disadvantages of their solutions. *Contemporary Educational Psychology*, *21*, 149–165. https://doi.org/10.1006/ceps.1996.0013

Borsboom, D., & Cramer, A. O. J. (2013). Network Analysis: An Integrative Approach to the Structure of Psychopathology. *Annual Review of Clinical Psychology*, *9*(1), 91–121. https://doi.org/10.1146/annurev-clinpsy-050212-185608

Brehm, J. W. (1999). The Intensity of Emotion. *Personality and Social Psychology Review*, *3*(1), 2–22. https://doi.org/10.1207/s15327957pspr0301_1

Brick, C., Hood, B., Ekroll, V., & de-Wit, L. (2022). Illusory Essences: A Bias Holding Back Theorizing in Psychological Science. *Perspectives on Psychological Science*, *17*(2), 491–506. https://doi.org/10.1177/1745691621991838

Byrne, B. M. (2001). Structural Equation Modeling With AMOS, EQS, and LISREL: Comparative Approaches to Testing for the Factorial Validity of a Measuring Instrument. *International Journal of Testing*, *1*(1), 55–86. https://doi.org/10.1207/S15327574IJT0101_4

Churchland, P. M. (1979). *Scientific Realism and the Plasticity of Mind*. Cambridge University Press. https://doi.org/10.1017/CBO9780511625435

Cogliati Dezza, I., Schulz, E., & Wu, C. M. (Eds.). (2022). *The Drive for Knowledge: The Science of Human Information Seeking*. Cambridge University Press. https://doi.org/10.1017/9781009026949

Corbit, L. H., & Balleine, B. W. (2005). Double Dissociation of Basolateral and Central Amygdala Lesions on the General and Outcome-Specific Forms of Pavlovian-Instrumental Transfer. *Journal of Neuroscience*, *25*(4), 962–970. https://doi.org/10.1523/JNEUROSCI.4507-04.2005

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281–302.

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Dalege, J., Borsboom, D., van Harreveld, F., van den Berg, H., Conner, M., & van der Maas, H. L. J. (2016). Toward a formalized account of attitudes: The Causal Attitude Network (CAN) model. *Psychological Review*, *123*(1), 2–22. https://doi.org/10.1037/a0039802

Dang, J., King, K. M., & Inzlicht, M. (2020). Why Are Self-Report and Behavioral Measures Weakly Correlated? *Trends in Cognitive Sciences*, *24*(4), 267–269. https://doi.org/10.1016/j.tics.2020.01.007

Danziger, K. (1990). *Constructing the Subject: Historical Origins of Psychological Research*. Cambridge University Press.

Dayan, P., & Abbott, L. F. (2005). *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems* (Revised ed. edition). The MIT Press.

Deci, E. (1992). The relation of interest to the motivation of behavior: A self-determination theory perspective. *The Role of Interest in Learning and Development*.

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*(6), 627.

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

Dennett, D. C. (1987). *The intentional stance* (pp. xi, 388). The MIT Press.

Descartes, R. (1955). *The philosophical works of Descartes. [2 vols.]* (pp. ix, 832). Dover Publications.

DeYoung, C. G., & Krueger, R. F. (2020). To Wish Impossible Things: On the Ontological Status of Latent Variables and the Prospects for Theory in Psychology. *Psychological Inquiry*, *31*(4), 289–296. https://doi.org/10.1080/1047840X.2020.1853462

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).

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Dobs, K., Martinez, J., Kell, A. J., & Kanwisher, N. (2022). Brain-like functional specialization emerges spontaneously in deep neural networks. *Science Advances*, *8*(11), eabl8913.

Donnellan, E., Usami, S., & Murayama, K. (2022). *Random item slope regression: Examining both similarities and differences in the association with individual items.* PsyArXiv. https://doi.org/10.31234/osf.io/s6erz

Dweck, C. S. (2017). From needs to goals and representations: Foundations for a unified theory of motivation, personality, and development. *Psychological Review*, *124*(6), 689–719. https://doi.org/10.1037/rev0000082

Eccles, J. S., & Wigfield, A. (2002). Motivational Beliefs, Values, and Goals. *Annual Review of Psychology*, *53*(1), 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153

Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, *61*, 101859. https://doi.org/10.1016/j.cedpsych.2020.101859

Eisenberg, I. W., Bissett, P. G., Zeynep Enkavi, A., Li, J., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Uncovering the structure of self-regulation through data-driven ontology discovery. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-10301-1

Elliot, A. J. (1997). Integrating 'classic' and 'contemporary' approaches to achievement motivation: A hierarchical model of approach and avoidance achievement motivation. In P. R. Pintrich & M. L. Maehr (Eds.), *Advances in motivation and achievement* (Vol. 10, pp. 143–179). JAI Press.

Elliot, A. J. (2023). Energization and Direction Are Both Essential Parts of Motivation. In M. Bong, J. Reeve, & S. Kim (Eds.), *Motivation Science: Controversies and Insights* (p. 0). Oxford University Press. https://doi.org/10.1093/oso/9780197662359.003.0002

Elliot, A. J., & Covington, M. V. (2001). Approach and avoidance motivation. *Educational Psychology Review*, *13*, 73–92.

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Elliot, A. J., & Fryer, J. W. (2008). The goal construct in psychology. In *Handbook of motivation science.* (pp. 235–250). New York, NY, US: Guilford Press.

Elliot, A. J., & Moller, A. C. (2003). Performance-approach goals: Good or bad forms of regulation? *International Journal of Educational Research*, *39*, 339–356.

Eronen, M. I. (2020). Causal discovery and the problem of psychological interventions. *New Ideas in Psychology*, *59*, 100785. https://doi.org/10.1016/j.newideapsych.2020.100785

Festinger, L. (1957). *A theory of cognitive dissonance.* Stanford University Press.

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

Fishbach, A., & Ferguson, M. J. (2007). The goal construct in social psychology. In *Social psychology: Handbook of basic principles, 2nd ed* (pp. 490–515). The Guilford Press.

Fiske, A. P. (2020). The lexical fallacy in emotion research: Mistaking vernacular words for psychological entities. *Psychological Review*, *127*(1), 95–113. https://doi.org/10.1037/rev0000174

FitzGibbon, L., Lau, J. K. L., & Murayama, K. (2020). The seductive lure of curiosity: Information as a motivationally salient reward. *Current Opinion in Behavioral Sciences*, *35*, 21–27. https://doi.org/10.1016/j.cobeha.2020.05.014

Fitzgibbon, L., & Murayama, K. (2022). Counterfactual curiosity: Motivated thinking about what might have been. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *377*(1866), 20210340. https://doi.org/10.1098/rstb.2021.0340

Frijda, N. H. (1986). *The emotions* (pp. xii, 544). Editions de la Maison des Sciences de l'Homme.

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1211–1221. https://doi.org/10.1098/rstb.2008.0300

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Gladwin, T. E., Figner, B., Crone, E. A., & Wiers, R. W. (2011). Addiction, adolescence, and the integration of control and motivation. *Developmental Cognitive Neuroscience*, *1*(4), 364–376. https://doi.org/10.1016/j.dcn.2011.06.008

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

Gottlieb, J., Oudeyer, P.-Y., Lopes, M., & Baranes, A. (2013). Information seeking, curiosity and attention: Computational and neural mechanisms. *Trends in Cognitive Sciences*, *17*(11), 585–593. https://doi.org/10.1016/j.tics.2013.09.001

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

Halliday, T. H. (1983). Motivation. In T. H. Halliday & P. J. B. Slater (Eds.), *Animal behaviour: Causes and effects* (Vol. 1, pp. 100–133). Oxford, UK: Blackwell Scientific Publications

Harackiewicz, J. M., & Sansone, C. (1991). Goals and intrinsic motivation: You can get there from here. In M. L. Maehr & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (Vol. 7, pp. 21–49). JAI Press.

Harlow, H. F. (1949). The formation of learning sets. *Psychological Review*, *56*(1), 51–65. https://doi.org/10.1037/h0062474

Heider, F. (1958). *The psychology of interpersonal relations* (pp. ix, 326). John Wiley & Sons Inc. https://doi.org/10.1037/10628-000

Hernán, M. A., & VanderWeele, T. J. (2011). Compound Treatments and Transportability of Causal Inference. *Epidemiology (Cambridge, Mass.)*, *22*(3), 368–377. https://doi.org/10.1097/EDE.0b013e3182109296

Hidi, S. E., & Renninger, K. A. (2019). Interest Development and Its Relation to Curiosity: Needed Neuroscientific Research. *Educational Psychology Review*, *31*(4), 833–852. https://doi.org/10.1007/s10648-019-09491-3

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Hinde, R. A. (1960). Energy models of motivation. *Symposia of the Society for Experimental Biology*, 14, 199–213.

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

Hulleman, C., Schrager, S., Bodmann, S., & Harackiewicz, J. (2010). A Meta-Analytic Review of Achievement Goal Measures: Different Labels for the Same Constructs or Different Constructs With Similar Labels? *Psychological Bulletin*, *136*, 422–449. https://doi.org/10.1037/a0018947

Jach, H. K., DeYoung, C. G., & Smillie, L. D. (2022). Why do people seek information? The role of personality traits and situation perception. *Journal of Experimental Psychology: General*, *151*, 934–959. https://doi.org/10.1037/xge0001109

James, W. (1890). *The Principles of Psychology: Vol. NA*.

Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, *34*(2), 183–202. https://doi.org/10.1007/BF02289343

Kanfer, R., & Chen, G. (2016). Motivation in organizational behavior: History, advances and prospects. *Organizational Behavior and Human Decision Processes*, *136*, 6–19. https://doi.org/10.1016/j.obhdp.2016.06.002

Kanwisher, N., Khosla, M., & Dobs, K. (2023). Using artificial neural networks to ask ‘why’ questions of minds and brains. *Trends in Neurosciences*, *46*(3), 240–254. https://doi.org/10.1016/j.tins.2022.12.008

Kelley, T. L. (1927). *Interpretation of educational measurements* (p. 353). World Book Co.

Kenny, A. (1971). The Homunculus Fallacy. In M. G. Grene & I. Prigogine (Eds.), *Interpretations of Life and Mind* (pp. 155–165). New York: Humanities Press.

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

Kihlstrom, J. F. (1987). The cognitive unconscious. *Science*, *237*(4821), 1445–1452. https://doi.org/10.1126/science.3629249

Kim, S.-I. (2013). Neuroscientific model of motivational process. *Front Psychol*, *4*, 98–98. https://doi.org/10.3389/fpsyg.2013.00098

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Kleinginna, P. R., & Kleinginna, A. M. (1981). A categorized list of motivation definitions, with a suggestion for a consensual definition. *Motivation and Emotion*, *5*(3), 263–291. https://doi.org/10.1007/BF00993889

Koch, S. (1941). The logical character of the motivation concept. I. *Psychological Review*, *48*, 15–38. https://doi.org/10.1037/h0062042

Koch, S. (1956). Behavior as" intrinsically" regulated: Work notes towards a pre-theory of phenomena called" motivational." In M. R. Jones (Ed.), *Nebraska Symposium on Motivation*, *4*, 42–87.

Kruglanski, A. W., Chernikova, M., & Kopetz, C. (2015). Motivation Science. In *Emerging Trends in the Social and Behavioral Sciences* (pp. 1–16). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118900772.etrds0104

Lau, H. C. (2009). Volition and the Function of Consciousness. In N. Murphy, G. F. R. Ellis, & T. O'Connor (Eds.), *Downward Causation and the Neurobiology of Free Will* (pp. 153–169). Springer. https://doi.org/10.1007/978-3-642-03205-9_9

Lazowski, R. A., & Hulleman, C. S. (2016). Motivation Interventions in Education: A Meta-Analytic Review. *Review of Educational Research*, *86*(2), 602–640. https://doi.org/10.3102/0034654315617832

LeDoux, J. E. (2014). Coming to terms with fear. *Proceedings of the National Academy of Sciences*, *111*(8), 2871–2878. https://doi.org/10.1073/pnas.1400335111

Lewin, K. (1942). Field Theory and Learning. *Teachers College Record*, *43*(10), 215–242. https://doi.org/10.1177/016146814204301006

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*, e1. https://doi.org/10.1017/S0140525X1900061X

Lieshout, L. L. F. van, Vandenbroucke, A. R. E., Müller, N. C. J., Cools, R., & Lange, F. P. de. (2018). Induction and Relief of Curiosity Elicit Parietal and Frontal Activity. *Journal of Neuroscience*, *38*(10), 2579–2588. https://doi.org/10.1523/JNEUROSCI.2816-17.2018

Locke, E. A. (2001). Self-set goals and self-efficacy as mediators of incentives and personality. In *Work motivation in the context of a globalizing economy* (pp. 13–26). Lawrence Erlbaum Associates Publishers.

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting & task performance*. Englewood Cliffs, NJ, US: Prentice-Hall, Inc.

Locke, E. A., & Latham, G. P. (2004). What Should we do About Motivation Theory? Six Recommendations for the Twenty-First Century. *Academy of Management Review*, *29*(3), 388–403. https://doi.org/10.5465/amr.2004.13670974

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

MacCorquodale, K., & Meehl, P. E. (1948). On a distinction between hypothetical constructs and intervening variables. *Psychological Review*, *55*(2), 95–107. https://doi.org/10.1037/h0056029

Madsen, K. B. (1974). *Modern theories of motivation*: A comparative metascientific study. Wiley.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* W. H. Freeman and Company.

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. https://doi.org/10.1037/xge0000140

Maslow, A. H. (1943). *A theory of human motivation. 50*(4), 370–396. https://doi.org/10.1037/h0054346

McClelland, D. C. (1987). *Human motivation*. New York, NY, US: Cambridge University Press.

McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. (1976). *The achievement motive*. Oxford, England: Irvington.

McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review*, *96*, 690–702.

McClure, K., Jacobucci, R., & Ammerman, B. A. (2021). *Are Items More Than Indicators? An examination of psychometric homogeneity, item-specific effects,*

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

and consequences for structural equation models. PsyArXiv. https://doi.org/10.31234/osf.io/n4mxv

McDougall, W. (1909). *An introduction to social psychology: Vol. NA*.

McDougall, W. (1926). *An Introduction to Social Psychology*. Methuen.

Murayama, K. (2022a). Are cognition, motivation, and emotion the same or different?: Let's abandon that thinking. In M. Bong, S. Kim, & J. Reeve (Eds.), *Motivation Science: Controversies and Insights.* Oxford University Press.

Murayama, K. (2022b). Motivation resides only in our language, not in our mental processes. In M. Bong, S. Kim, & J. Reeve (Eds.), *Motivation Science: Controversies and Insights.* Oxford University Press.

Murayama, K. (2022c). A reward-learning framework of knowledge acquisition: An integrated account of curiosity, interest, and intrinsic–extrinsic rewards. *Psychological Review*, *129*(1), 175–198. https://doi.org/10.1037/rev0000349

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., Izuma, K., Aoki, R., & Matsumoto, K. (2017). "Your Choice" Motivates You in the Brain: The Emergence of Autonomy Neuroscience. In S. Kim, J. Reeve, & M. Bong (Eds.), *Advances in Motivation and Achievement (Vol. 19, Recent Developments in Neuroscience Research on Human Motivation)* (pp. 95–125). Emerald.

Murphy, P. K., & Alexander, P. A. (2000). A Motivated Exploration of Motivation Terminology. *Contemporary Educational Psychology*, *25*(1), 3–53. https://doi.org/10.1006/ceps.1999.1019

Murray, H. A. (1938). *Explorations in personality.* Oxford University Press.

Nagengast, B., & Trautwein, U. (2023). Theoretical and Methodological Disintegration Is the Most Fundamental Limitation in Contemporary Motivation Research. In M. Bong, J. Reeve, & S. Kim (Eds.), *Motivation Science: Controversies and Insights* (p. 0). Oxford University Press. https://doi.org/10.1093/oso/9780197662359.003.0068

Newell, A. (1994). *Unified Theories of Cognition:* Harvard University Press.

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Ningombam, D. D., Yoo, B., Kim, H. W., Song, H. J., & Yi, S. (2022). CuMARL: Curiosity-Based Learning in Multiagent Reinforcement Learning. *IEEE Access*, *10*, 87254–87265. https://doi.org/10.1109/ACCESS.2022.3198981

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231–259. https://doi.org/10.1037/0033-295X.84.3.231

Niv, Y., Joel, D., & Dayan, P. (2006). A normative perspective on motivation. *Trends in Cognitive Sciences*, *10*, 375–381.

O'Reilly, R. C. (2020). Unraveling the Mysteries of Motivation. *Trends in Cognitive Sciences*, *24*(6), 425–434. https://doi.org/10.1016/j.tics.2020.03.001

Oudeyer, P.-Y. K., Frédéric. (2007). *What is Intrinsic Motivation? A Typology of Computational Approaches. 1*(NA), 6–6. https://doi.org/10.3389/neuro.12.006.2007

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

Palminteri, S., & Lebreton, M. (2022). The computational roots of positivity and confirmation biases in reinforcement learning. *Trends in Cognitive Sciences*, *26*(7), 607–621. https://doi.org/10.1016/j.tics.2022.04.005

Patankar, S. P., Zhou, D., Lynn, C. W., Kim, J. Z., Ouellet, M., Ju, H., Zurn, P., Lydon-Staley, D. M., & Bassett, D. S. (2022). Curiosity as filling, compressing, and reconfiguring knowledge networks. *ArXiv:2204.01182 [q-Bio]*. http://arxiv.org/abs/2204.01182

Pekrun, R. (2023). Jingle-Jangle Fallacies in Motivation Science: Toward a Definition of Core Motivation. In M. Bong, J. Reeve, & S. Kim (Eds.), *Motivation Science: Controversies and Insights* (p. 0). Oxford University Press. https://doi.org/10.1093/oso/9780197662359.003.0009

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Pekrun, R., & Marsh, H. W. (2022). Research on situated motivation and emotion: Progress and open problems. *Learning and Instruction*, *81*, 101664. https://doi.org/10.1016/j.learninstruc.2022.101664

Pessoa, L., Medina, L., & Desfilis, E. (2022). Refocusing neuroscience: Moving away from mental categories and towards complex behaviours. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *377*(1844), 20200534. https://doi.org/10.1098/rstb.2020.0534

Pittman, T. S., & Zeigler, K. R. (2007). Basic human needs. In *Social psychology: Handbook of basic principles, 2nd ed* (pp. 473–489). The Guilford Press.

Poli, F., Meyer, M., Mars, R. B., & Hunnius, S. (2022). Contributions of expected learning progress and perceptual novelty to curiosity-driven exploration. *Cognition*, *225*, 105119. https://doi.org/10.1016/j.cognition.2022.105119

Reeder, G. (2009). Mindreading: Judgments About Intentionality and Motives in Dispositional Inference. *Psychological Inquiry*, *20*(1), 1–18. https://doi.org/10.1080/10478400802615744

Reeve, J. (2017). *Understanding motivation and emotion (7th Ed.)*. Wiley.

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). Academic Press.

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

Rhemtulla, M., van Bork, R., & Borsboom, D. (2020). Worse than measurement error: Consequences of inappropriate latent variable measurement models. *Psychological Methods*, *25*(1), 30–45. https://doi.org/10.1037/met0000220

Richter, M., Gendolla, G. H. E., & Wright, R. A. (2016). Three decades of research on motivational intensity theory: What we have learned about effort and what we still don't know. *Advances in Motivation Science*, *3*, 149–186.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill, factors predict college outcomes? A meta-analysis.

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Psychological Bulletin, *130*(2), 261–288. https://doi.org/10.1037/0033-2909.130.2.261

Rohrer, J. M., & Murayama, K. (2021). *These are not the effects you are looking for: Causality and the within-/between-person distinction in longitudinal data analysis*. PsyArXiv. https://doi.org/10.31234/osf.io/tg4vj

Rolls, E. T. (2016). Motivation Explained: Ultimate and Proximate Accounts of Hunger and Appetite. In A. J. Elliot (Ed.), *Advances in Motivation Science* (Vol. 3, pp. 187–249). Elsevier. https://doi.org/10.1016/bs.adms.2015.12.004

Rozek, C. S., Svoboda, R. C., Harackiewicz, J. M., Hulleman, C. S., & Hyde, J. S. (2017). Utility-value intervention with parents increases students' STEM preparation and career pursuit. *Proceedings of the National Academy of Sciences*, *114*(5), 909–914. https://doi.org/10.1073/pnas.1607386114

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. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. 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

Sachisthal, M. S. M., Jansen, B. R. J., Peetsma, T. T. D., Dalege, J., van der Maas, H. L. J., & Raijmakers, M. E. J. (2019). Introducing a science interest network model to reveal country differences. *Journal of Educational Psychology*, *111*(6), 1063–1080. https://doi.org/10.1037/edu0000327

Schultheiss, O. C., Campbell, K. L., & McClelland, D. C. (1999). Implicit Power Motivation Moderates Men's Testosterone Responses to Imagined and Real Dominance Success. *Hormones and Behavior*, *36*(3), 234–241. https://doi.org/10.1006/hbeh.1999.1542

Schultheiss, O. C., Yankova, D., Dirlikov, B., & Schad, D. J. (2009). Are Implicit and Explicit Motive Measures Statistically Independent? A Fair and Balanced Test

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Using the Picture Story Exercise and a Cue- and Response-Matched Questionnaire Measure. *Journal of Personality Assessment*, *91*(1), 72–81. https://doi.org/10.1080/00223890802484456

Sedikides, C., & Strube, M. J. (1997). Self-evaluation: To thine own self be good, to thine own self be sure, to thine own self be better. *Advances in Experimental Social Psychology*, *29*, 209–269.

Seward, G. H. (1939). Dialectic in the psychology of motivation. *Psychological Review*, *46*, 46–61. https://doi.org/10.1037/h0057732

Sheldon, K. M. (2022). Freely Determined: What the New Psychology of the Self Teaches Us About How to Live. Basic Books.

Shultz, T. R., & Lepper, M. R. (1996). Cognitive dissonance reduction as constraint satisfaction. *Psychological Review*, *103*(2), 219–240. https://doi.org/10.1037/0033-295X.103.2.219

Simpson, E. H., & Balsam, P. D. (2016). The Behavioral Neuroscience of Motivation: An Overview of Concepts, Measures, and Translational Applications. *Current Topics in Behavioral Neurosciences*, *27*, 1–12. https://doi.org/10.1007/7854_2015_402

Skinner, E. A. (1996). A guide to constructs of control. Journal of Personality and Social Psychology, 71, 549–570.

Smith, B. J., & Read, S. J. (2022). Modeling incentive salience in Pavlovian learning more parsimoniously using a multiple attribute model. *Cognitive, Affective, & Behavioral Neuroscience*, *22*(2), 244–257. https://doi.org/10.3758/s13415-021-00953-2

Smith, G. T., Fischer, S., & Fister, S. M. (2003). Incremental Validity Principles in Test Construction. *Psychological Assessment*, *15*(4), 467–477.

Steele, C. M. (1988). The Psychology of Self-Affirmation: Sustaining the Integrity of the Self. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 21, pp. 261–302). Academic Press. https://doi.org/10.1016/S0065-2601(08)60229-4

Steinman, M. Q., Duque-Wilckens, N., & Trainor, B. C. (2019). Complementary neural circuits for divergent effects of oxytocin: Social approach versus social anxiety.

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Biological Psychiatry, *85*(10), 792–801. https://doi.org/10.1016/j.biopsych.2018.10.008

Stevens, L. E., & Fiske, S. T. (1995). Motivation and Cognition in Social Life: A Social Survival Perspective. *Social Cognition*, *13*(3), 189–214. https://doi.org/10.1521/soco.1995.13.3.189

Stich, S. P., & Ravenscroft, R. (1994). What is folk psychology? *Cognition*, *50*, 447–468.

Tamura, A., Ishii, R., Yagi, A., Fukuzumi, N., Hatano, A., Sakaki, M., Tanaka, A., & Murayama, K. (2022). Exploring the within-person contemporaneous network of motivational engagement. *Learning and Instruction*, *81*, 101649. https://doi.org/10.1016/j.learninstruc.2022.101649

Utman, C. H. (1997). Performance Effects of Motivational State: A Meta-Analysis. *Personality and Social Psychology Review*, *1*(2), 170–182. https://doi.org/10.1207/s15327957pspr0102_4

van der Maas, H. L. J., Dolan, C. V., Grasman, R. P. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review*, *113*(4), 842–861. https://doi.org/10.1037/0033-295X.113.4.842

van Rooij, I., & Baggio, G. (2021). Theory Before the Test: How to Build High-Verisimilitude Explanatory Theories in Psychological Science. *Perspectives on Psychological Science*, *16*(4), 682–697. https://doi.org/10.1177/1745691620970604

Vancouver, J. B., Wang, M., & Li, X. (2020). Translating Informal Theories Into Formal Theories: The Case of the Dynamic Computational Model of the Integrated Model of Work Motivation. *Organizational Research Methods*, *23*(2), 238–274. https://doi.org/10.1177/1094428118780308

VandenBos, G. R., & American Psychological Association (Eds.). (2007). *APA dictionary of psychology* (1st ed). American Psychological Association.

VanderWeele, T. J. (2022). Constructed Measures and Causal Inference: Towards a New Model of Measurement for Psychosocial Constructs. *Epidemiology*, *33*(1), 141. https://doi.org/10.1097/EDE.0000000000001434

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Wegner, D. M. (2004). Précis of The illusion of conscious will. *Behavioral and Brain Sciences*, *27*(5), 649–659. https://doi.org/10.1017/S0140525X04000159

Weiner, B. (1992). *Human motivation: Metaphors, theories, and research.* Sage Publications.

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

Wigfield, A., Muenks, K., & Eccles, J. S. (2021). Achievement Motivation: What We Know and Where We Are Going. *Annual Review of Developmental Psychology*, *3*(1), 87–111. https://doi.org/10.1146/annurev-devpsych-050720-103500

Zhang, J., Berridge, K. C., Tindell, A. J., Smith, K. S., & Aldridge, J. W. (2009). A Neural Computational Model of Incentive Salience. *PLOS Computational Biology*, *5*(7), e1000437. https://doi.org/10.1371/journal.pcbi.1000437

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Figure legends

Figure 1

Motivation as the (black box) determinant of behavior (A) and motivation as psychological construction of underlying mental computational processes (B). Graphics are available under creative commons licenses (https://creativecommons.org/), downloaded from https://upload.wikimedia.org/wikipedia/commons/d/d9/A_black_box.svg, https://commons.wikimedia.org/wiki/File:Gear_-_Noun_project_7137.svg and https://commons.wikimedia.org/wiki/File:Gears-686316.jpg

Figure 2

Reward-learning framework of knowledge acquisition (Murayama, 2022).

Figure 3

Different interpretations of statistical results according to different perspectives.

Figure 4.

Direction aspect of motivation: Motivation functions as utility mapping.

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

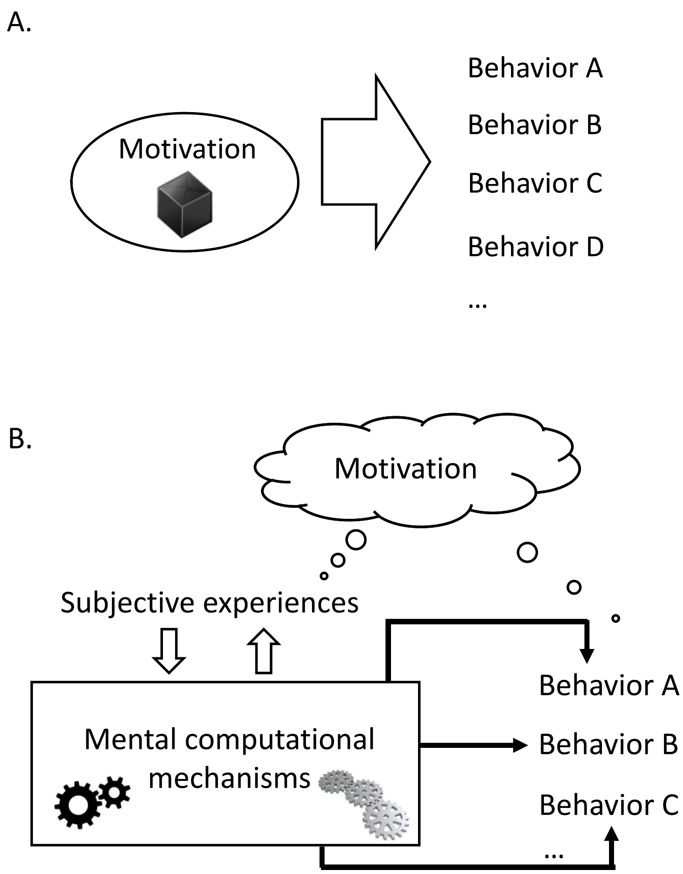


Figure 1

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

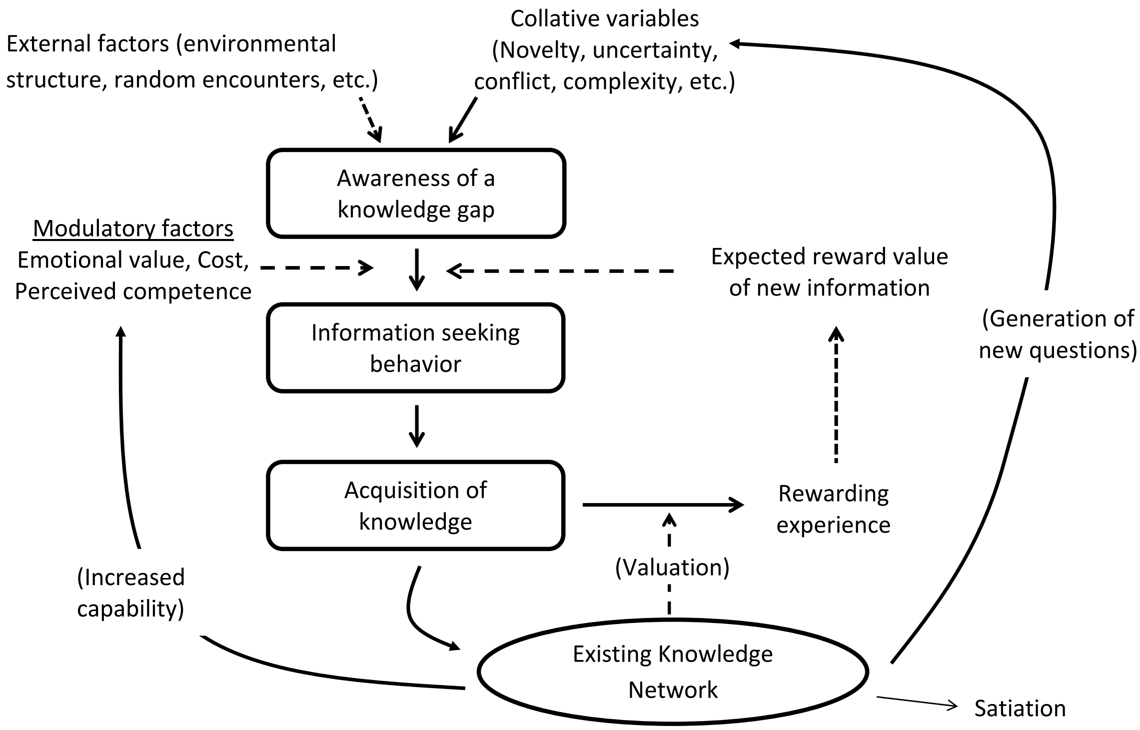


Figure 2

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

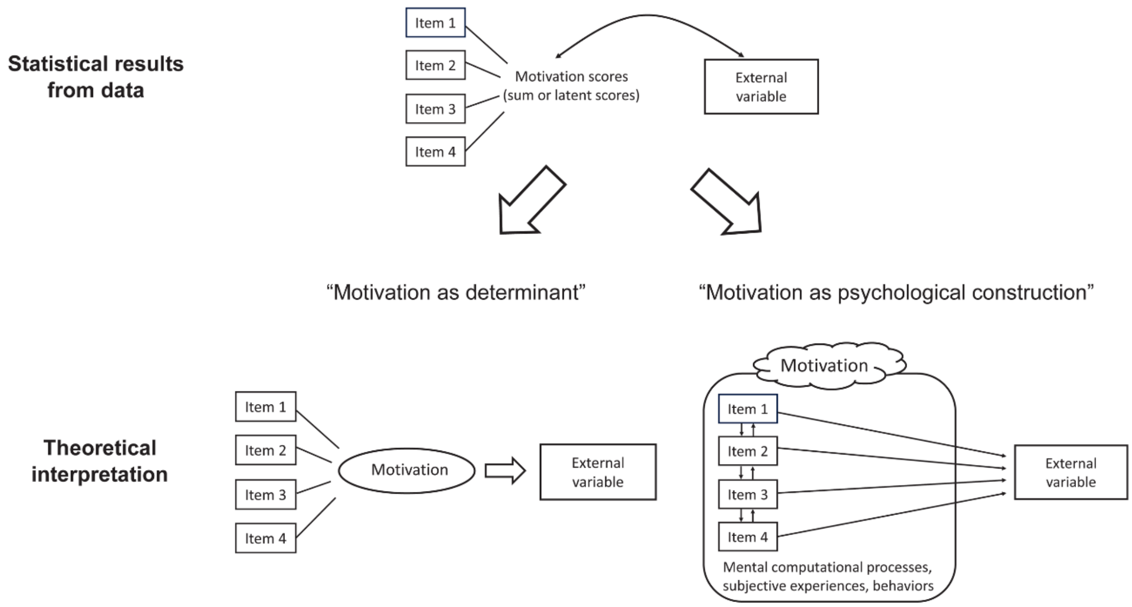
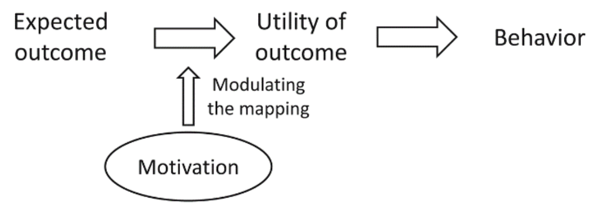


Figure 3.

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press



https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press

Figure 4.

https://doi.org/10.1017/S0140525X24000025 Published online by Cambridge University Press