TOPICS
TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science 00 (2026) e70082
© 2026 Cognitive Science Society LLC
ISSN: 1756-8765 online
DOI: 10.1111/tops.70082

This article is part of the topic "Goal Dynamics in Cognition," Gaia Molinaro, Cédric Colas, Junyi Chu, and Robert Hawkins (Topic Editors).

The Syntax and Semantics of Goals

David M. Abel,[a] Mark K. Ho[b]

[a]*Department of Informatics, University of Edinburgh*
[b]*Department of Psychology, New York University*

Received 1 November 2025; received in revised form 20 July 2026; accepted 28 July 2026

Abstract

In both cognitive science and computer science, goals are conceptualized as cognitive states that flexibly combine with world knowledge to organize and specify purposeful behavior. In this way, *goals are compositional representations whose content relates to rational behavior*. We here draw attention to goals as representations and their content because it highlights a parallel with other areas in cognitive science—in particular, the syntax–semantics interface in linguistics and logic—while also foregrounding foundational questions about the *expressivity*, *design*, and *efficiency* of different goal representations. For example, goals are typically taken as fixed and imposing constraints on desirable behaviors, but we can also identify constraints on goal representations themselves, such as whether a particular goal language is sufficiently expressive to capture behaviors of interest, or whether different goal representations capture the same behavior. Here, we synthesize work that aims to characterize the properties of different goal representations and suggest these are points of a broader design space. We close by discussing how distinguishing the form and meaning of goals can elucidate the implicit assumptions we make about goals, inform the study of interactions between higher-level cognition and motivation, and isolate axes of variation for different conceptions of goals.

Keywords: Goals; Reward; Preference; Semantics; Syntax; Value

Correspondence should be sent to David M. Abel, Department of Informatics, University of Edinburgh, 10 Crichton Street, Edinburgh EH8 9AB, UK. E-mail: david.abel@ed.ac.uk
All rights reserved, including rights for text and data mining and training of artificial intelligence technologies or similar technologies.

1. Introduction

Goal representations are central to the study of biological intelligences (Ball, 2023) and the design of artificial intelligences (Russell & Norvig, 2010). Indeed, our intuitions about goals underpin much of our everyday theory of mind (Premack & Woodruff, 1978). But, what are goals? What distinguishes goals from other psychological states, such as reflexes, habits, or knowledge about the world? How do goals interact with these other representations? And if goals are a kind of representation, what exactly do they represent?

To address these questions, we begin by highlighting two key properties of goal representations: first, *goals control rational behavior*; second, *goals are compositional*. The first property, that goals control rational behavior, simply means that goals are control representations that relate to behavior via a principle of rationality (Cohen, 2017; Dennett, 1989; Miller & Cohen, 2001). As an example, consider the goal to *run the Boston marathon*. To successfully run the Boston marathon, one can do several things to make that more likely, including: devise a training regimen, eat healthy, not skip practice, and show up at the starting line on the day of the marathon. Having a goal like running the Boston marathon is to have a mental state that orients behaviors toward those likely to achieve the goal and away from those unlikely to achieve the goal; in other words, goals orient an intelligent agent's behavior toward instrumentally rational, teleological action (Bratman, 1987; Dennett, 1989; Dickinson & Balleine, 2000). This capacity to rationally control or override otherwise nonrational behavior distinguishes goals from other representations, such as desires, which may not affect behavior, or habits, which may not be rational (Berridge & Robinson, 2003; Wood & Rünger, 2016). For this reason, goals are viewed as central to processes that override or flexibly configure behavior (or even other cognitive processes) in a rational, task-dependent manner. This includes processes such as planning (Mattar & Lengyel, 2022), meta-cognition (Griffiths et al., 2019), social cognition (Jara-Ettinger et al., 2016), self-regulation (Latham & Locke, 1991), cognitive control (Cohen, 2017), and mental effort (Shenhav et al., 2017).

The second property, that *goals are compositional*, means that goal representations combine with other representations in a manner that systematically relates to behavior (Bakermans, Warren, Whittington, & Behrens, 2025; Davidson & Gureckis, 2024; Velez-Ginorio, Siegel, Tenenbaum, & Jara-Ettinger, 2017). For example, given two simple goals, *run the Boston marathon* and *avoid injury*, one can combine them into a composite goal, *run the Boston marathon without getting injured*. This new goal specifies behavior in ways that are systematically similar to and different from those specified by the original goal—for example, one will still go for practice runs but only gradually increase the length of practice runs to avoid injury. Moreover, not only do goals compose with other goals, they flexibly compose with new knowledge about the world. For instance, if one wanted to avoid injury and then learned about the importance of stretching consistently for that goal, one would change their behavior to incorporate stretching into their practice routine. Indeed, the property that goals can compose with world knowledge to flexibly guide behavior is perhaps *the* hallmark of goal-directed behavior. Experimental manipulations such as *reward devaluation*, in which a previously rewarding outcome stops being rewarding (e.g., moving water in a maze to a new location), explicitly leverage this property to assess whether a behavior is goal-directed

17568765, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tops.70082 by Mark Ho - NYU Medical Central Library , Wiley Online Library on [27/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

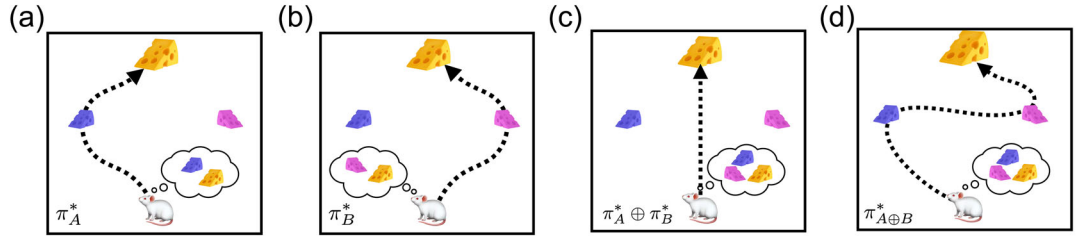


Fig. 1. Behavioral composition is not rationality-preserving. A rat could have the goal of (A) reaching the big cheese and the small blue cheese or (B) the big cheese and the small pink cheese. The optimal behaviors are represented as dotted lines and denoted π^*_A and π^*_B, respectively. (C) Directly composing the *behaviors* from A and B (e.g., $\pi^*_A \oplus \pi^*_B$) might have the rat attempt to walk straight to the cheese. (D) This is different from first composing the *goals* of A and B to derive a compound goal—reach the big cheese and both small colored cheeses—and then acting rationally with respect to the compound goal (i.e., $\pi^*_A \oplus \pi^*_B \neq \pi^*_{A \oplus B}$).

or simply habitual (e.g., if a rat stops navigating to the former location of the water given minimal training, this is taken as evidence for a goal representation) (Dickinson, 1985).

In short, goals control rational behavior and goals are compositional.[1] However, we argue that the precise way in which goals and behaviors compose and relate to one another is far from obvious. Consider the example shown in Fig. 1, in which a rat could be in two situations: one in which they have the goal of reaching the big yellow cheese and the small blue cheese (goal A), and, separately, in which they have the goal of reaching the big yellow cheese and the small pink cheese (goal B). A and B each produce different goal-directed behaviors, which we can call behaviors A and B. Notice, however, that while we could directly compose behaviors A and B by, say, averaging (panel C), this is not the same as the rational behavior for the composed goal $A \oplus B$—that is, reach the cheese while avoiding *both* obstacles (panel D). This example illustrates a general problem: Even though goals control rational behavior and are compositional, *direct composition of rational behavior does not preserve rationality*.

1.1. The syntax–semantics interface

The preceding discussion highlights a tension between goals as specifications of rational behavior and goals as representations that systematically compose. We will not fully resolve this tension in this paper. Rather, our aim is to provide a lens for clarifying the core issues and frame recent research from this perspective. To this end, our primary claim is that the study of goals should borrow an analytical distinction developed in another area of cognitive science: the syntax–semantics interface from linguistics and logic (Jacobson, 2014; Montague, 1970).

In the study of natural and artificial languages, semantics focuses on the *meaning* of expressions. For example, the word "Rex" can refer to a dog, while the phrase "the neighbor's cat" can refer to a particular cat. In contrast, syntax focuses on how simpler linguistic elements combine to form complex expressions and the rules for combining them, also called a *grammar*. For example, in English, "Rex chased the neighbor's cat" is a well-formed

17568765, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tops.70082 by Mark Ho - NYU Medical Central Library , Wiley Online Library on [27/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

expression consisting of the words "Rex," "chased," "the," "neighbor's," and "cat," whereas "chased Rex the neighbor's cat" is not. The syntax–semantics interface concerns how the form of expressions interacts with the meaning of expressions. In particular, given a language, one can ask about its *compositional semantics*, that is, how the semantic meaning of complex expressions derives from the meaning of its simpler constituents and the syntactic rules used to combine them (Jacobson, 2014). The utterance "Rex chased the neighbor's cat," for instance, denotes a state of affairs in which a dog named Rex chased a cat, and its meaning is derived from the meanings of its constituents (i.e., "Rex," "chased," and "the neighbor's cat") and their composition (i.e., a subject-verb-object construction).

A complete review of the syntax–semantics distinction and its history is beyond our scope, but it is worth outlining some ways in which it serves as a valuable analytic tool in the study of language. First, the distinction between the syntax and semantics of a language allows us to talk precisely about *semantic equivalence*. For example, in English, "Rex chased the neighbor's cat" means effectively the same thing as "The neighbor's cat was chased by Rex" because they denote the same state of affairs. The difference between these two statements is their syntax: the first statement uses an active construction, while the second statement uses a passive construction (Siewierska, 1984). Second, we can ask about the *expressivity* of the lexicon and syntax of a language with respect to a semantics. Different languages provide different grammatical forms that can make it easier or harder to express certain meanings. For example, even though all languages can express the source of information, some language families (e.g., Turkic languages) have a distinct grammatical form for denoting how some fact was acquired (e.g., by direct experience or from another's testimony) that must always be specified (Johanson, 2003). Finally, a compositional semantics provides a framework for characterizing systematic patterns of *learning*. As an example, research on early word learning suggests that infants acquire nouns before verbs because nouns tend to denote objects, whereas verbs tend to denote actions and relations among objects (Golinkoff & Hirsh-Pasek, 2008). Nouns and verbs are syntactic notions, while objects and relations between them are semantic ones.

Beyond linguistics, the relationship between the syntactic and semantic features of a representational system is a recurring theme throughout cognitive science. For instance, compositionality has been argued to be an important general principle not only of language, but cognition and learning more broadly (Fodor & Pylyshyn, 1988; Kirby et al., 2015; Lake et al., 2017). Similarly, within computer science, delineating interactions between syntax and semantics is central to fields such as logic, programming languages, compilers, and formal verification (Pierce, 2002). Here, we propose that the syntax-semantic distinction can serve a similar role in organizing how we think about goals and agency.

2. The syntax and semantics of goal representations

How might the syntax–semantics distinction be applied to goals? Drawing on the analogy with language, any compositional semantics for goals needs to answer two questions: What is the semantic space being represented by goals? And what are the syntactic rules for composing goal representations?

17568765, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tops.70082 by Mark Ho - NYU Medical Central Library , Wiley Online Library on [27/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

We expect there to be many reasonable answers to these two questions that yield different implications. To ground our discussion, we start by focusing on how the syntax–semantics distinction might clarify issues in reinforcement learning, a formal framework for studying goal-directed behavior that is used in both cognitive science and artificial intelligence (Dayan & Niv, 2008; Sutton & Barto, 2018). After taking a deep dive into this one "point" in the "space" of possible compositional semantics, we will discuss how this perspective can shed light on possible goal languages and other representations that interface with goals.

2.1. The expressivity of goals in reinforcement learning

In standard formulations of reinforcement learning, tasks are most often characterized as Markov decision processes (MDPs: Puterman, 2014). The commitment to the MDP as a mechanism for modeling tasks forefronts the two syntax–semantics questions directly: what are the semantics of goals in an MDP, and what are the syntactic rules for composing goals? Within an MDP, an agent's goal is represented by a reward function that assigns a number to each possible experience. As articulated by the reward hypothesis (Sutton, 2004), the agent's goal is then understood in terms of maximizing some accumulation of reward (typically, an expected discounted sum).

In this way, we have a clear syntactic picture: goals are *represented* as rewards—numbers assigned to each experience—which can combine across time and possible consequences to form a composite value. However, this still does not provide a *semantics* to these goals: what does a reward of "7" mean? What does the sequence of rewards "2, –1, 0, 4" mean? In other words, if we have two agents with two separate goals, how can we answer questions about the meaning of their respective reward sequences, or their reward functions more generally? To do so, we suggest it is necessary to interpret reward as a syntactic structure whose meaning is grounded in some other content—in this case, preferences over behavior.

In recent work, Abel et al. (2021) make this study explicit by asking: what can reward functions of different kinds *express*? In this way, they draw a relationship between the syntactic form of reward (its numerical value) and the meaning referred to by these values. Specifically, the arguments presented by Abel et al. (2021) suggest that *reward* is a communicative mechanism for the representational content the reward refers to: a preference relation on policies. That is, Abel et al. consider (among several choices) a simple class of preference relations over policies they call a Set Of Acceptable Policies (SOAPs). A SOAP is a preference relation on the class of deterministic functions that assign action to state. By definition, a SOAP only induces two equivalence classes: the good strategies for acting, and the bad. Abel et al. ask: Can Markov reward functions *express* each possible SOAP? A Markov reward function is so-called as it only depends on the most recent state experienced, as opposed to the full history of interaction. Implicitly, this question is one that exploits the syntax and semantics of goals—can reward, a syntactic structure, contain the necessary representational content of a particular kind of preference ordering? Thus, it is a question about the expressive power of the *syntactic language of reward* for referring to the *semantic language of preferences*.

Crucially, without further commitments, a preference relation under-determines a reward function. It is similar in spirit to a set of logical predicates without their connectives, or a

17568765, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tops.70082 by Mark Ho - NYU Medical Central Library , Wiley Online Library on [27/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

choice of which logical rules to embrace: we have not yet made a commitment for *how* the reward needs to *compose* in order to form a value function that actually does the ordering. In logic, we might reject the law of excluded middle as famously argued by Priest (1983), or *Modus Ponens* as explored by Lewis (1895) and Decker (2006), or choose to embrace real-numbered truth values as in fuzzy logic (Zadeh, 1988). Such choices can impact the way we can compose our logical primitives into sentences, or the truth value of those corresponding sentences. In an MDP, even value is not semantic on its own; assigning a specific number to some state of affairs, such as "4.3," carries no extrinsic meaning. This is in part due to a key aspect of the decision theoretic foundations that underlie much of consequentialist views of decision-making. Specifically, the classical work by von Neumann and Morgenstern (vNM: von Neumann & Morgenstern, 1953) establishes that any utility function u that captures a vNM-satisfying preference is *unique up to affine transformations*. In this way, we could select two distinct utility functions u and u' that realize the exact same preference, but assign arbitrarily different utilities to an individual state of affairs. As a result, the precise number associated with value only takes on meaning in so far as it refers to a preference relation. Two utility functions can thus *mean* the same thing when they refer to the same preference ordering. The same argument holds of reward, thus solidifying their status as syntactic aspects of goals.

In goal space, Abel et al. (2021) explore this perspective in two steps. First, they turn reward into the standard compositional view of *expected cumulative discounted reward*, for a choice of geometric discount factor. A value function orders the states of the Markov process underlying the MDP. To connect the preferences implicit in SOAPs with these value functions, Abel et al. again take a standard view to treat the *expected cumulative discounted reward* of each specific policy in the SOAP, evaluated at the start-state of the MDP. Then, a reward function can *express* a SOAP just when the start-state value of every good policy is strictly higher than the start-state value of every bad policy. In other words: if a SOAP is a splitting of the decision-making strategies into "good" and "bad," can the start-state value function induced by a reward function define the boundary separating good from bad? Together, these commitments yield the following result about the expressivity of Markov reward.

Theorem 1 Informal, Theorem 4.1 of Abel et al. (2021). *For any finite Markov controlled process, there exists a SOAP such that there is no Markov reward function that expresses the SOAP.*

This result reveals a limitation to the expressivity of certain syntactic goal representations: Markov rewards.

Notably, expressivity is just one kind of question we can ask about the syntax of a goal. Indeed, much of the history on rewards has little to do with its expressivity compared to its computational affordances. Markov rewards are amenable to dynamic programming (Bellman, 1957), and, therefore, fit tidily into a wide variety of existing algorithms for both planning and learning. We suggest that a key set of questions for further research will determine the space of considerations of goals that live on the syntactic and semantic sides of the divide.

17568765, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tops.70082 by Mark Ho - NYU Medical Central Library , Wiley Online Library on [27/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.1.1. Settling the reward hypothesis

In Abel et al., the bridge from reward to preference come about largely due to the Markov property, when we move beyond the Markov property, the design space of such decisions becomes much richer. In subsequent work, Bowling, Martin, Abel, and Dabney (2023) expand on the analysis of Abel et al. (2021) in two ways. First, they go beyond Markovian environments and policies and instead consider the space of partially observable environments (Abel et al., 2023; Dong, Van Roy, & Zhou, 2022; Hutter, 2000; Kaelbling, Littman, & Cassandra, 1998) and history-based policies. In this case, the agent is only ever exposed to a stream of observations. As such, when we conceive of what the agent wants, these "wants" must manifest in terms of steering the agent's observable experience.

Second, Bowling et al. seek a complete characterization of the implicit requirements under which we can relate rewards and preferences. As with the previous work, this analysis demands both compositional and representational commitments. Here, the compositional commitments are again roughly made in terms of a standard recipe: expected, cumulative, discounted rewards are embraced as the default form of value (though discount is allowed to vary on a per-experience basis, as per White, 2017).

On the representational side, Bowling et al. provide two assumptions that create a complete bridge from reward, to value, to the preferences of interest. First, they assume that *what it means to prefer a policy* π_1 to another π_2 is that, eventually, we always prefer the outcomes or states of affairs produced by π_1 to those of π_2. This allows intermediate preferences over the outcomes induced by π_1 and π_2 to oscillate, possibly for an extended period, so long as they eventually settle into π_1's indefinite preference over π_2. This is one natural view of what it means to prefer one policy to another, but it is again just one choice. Second, they assume that a reward function expresses a preference over policies if the value function formed by that reward orders *outcomes* according to the policy preferences. As such, this pair of assumptions tells us both how to connect policy preferences with outcome preferences, and how to connect value to preference.

Then, subject to these two assumptions, Bowling et al. prove the following.

Theorem 2 Informal restatement of Theorem 4.1 by Bowling et al. (2023). *A Markov reward function is capable of representing preference over policies if and only if the preference satisfies the four vNM axioms and a fifth called Temporal* γ*-indifference.*

We refer the interested reader to the paper and proof by Bowling for full details on the result, and the axioms. For our purposes, we highlight several salient facts. First, we note that questions of expressivity are a meaningful way to understand the richness of different goal languages. Second, the semantic axioms come in the form of Assumptions 1 and 2 by Bowling et al. (2023), which provide a story for connecting value and preference (roughly: policy preferences correspond to finite-horizon outcome preferences). However, we suspect this is just the tip of the iceberg. Ávila Pires et al. (2025), for instance, expand on this account and explore variations of distributional reinforcement learning (Bellemare, Dabney, & Munos, 2017) based on *stock augmented* decision-making that facilitates goals of the kind "keep the

17568765, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tops.70082 by Mark Ho - NYU Medical Central Library , Wiley Online Library on [27/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

return of this cumulant within this range." A full exploration of possible semantic axioms represents an important direction for future work.

Moreover, while we have here focused our discussion on the particulars of the goal formalisms of reinforcement learning, we suggest that the basic template extends to all goal accounts: that is, goals, by necessity, adhere to the syntax-semantic framing.

2.2. Semantic expressivity of different goal languages

The work reviewed above primarily focuses on the syntax and semantics of goals from the perspective of standard reinforcement learning, with rewards, values, and preferences as the substrate of goals. A complete review of the literature that explores these issues is beyond our scope, but we note a few key lines of work that address questions about goal syntax and semantics. Davidson, Todd, Togelius, Gureckis, and Lake (2025), for instance, ground goals in terms of *reward-producing programs*, a view that is both resonant with Sutton's reward hypothesis, but also adds further structure to how these goals are represented. Other approaches include purely logical forms of goals, as in classical planning (Newell et al., 1959) that typically amount to a predicate on states of affairs. Variations of this kind add further richness through incorporating temporal operators, such as linear temporal logic (Baier & McIlraith, 2006; Hammond et al., 2021; Littman et al., 2017; Richens et al., 2025), and the reward machines they are often represented by Icarte, Klassen, Valenzano, and McIlraith (2022). Others further incorporate elements of *risk-sensitivity* explicitly into goals (Howard & Matheson, 1972; Lehman et al., 2025; Mihatsch & Neuneier, 2002). Risk-sensitive views are siblings of a broader class of goal accounts that are based in multiple objectives (Gábor et al., 1998)—that is, a decision-maker might have many qualities they care about, and the pursuit of goals is ultimately about balancing these quantities. For example, Ringstrom (2023) suggests that thirst and hunger might be fundamental, atomic pursuits for decision-makers. Further accounts reject the presence of a discount factor entirely and embrace a truly infinite horizon using the *average reward* (Mahadevan 1996; Dong et al. 2022; Kumar et al. 2023). Other proposals specifically build around roots in virtue ethics, rather than consequentialism, as articulated by Chang (2015). Here, Chang argues against the notions of being able to compare any two experiences, as echoed by Aumann (1962) in his argument against one of the vNM axioms.

As articulated by Davidson and Gureckis (2024), different goal representations offer different affordances, even if they ultimately represent the same thing. For instance, two different representations of the same goal can afford different types or degrees of compositionality, generalization, and learning efficiency. For example, Tasse et al. (2020) and Tasse (2020) develop a boolean algebra for *logically* composing goals. Subject to a few minor assumptions, rewards of the usual kind can be cast as members of a boolean algebra, unlocking the operations of conjunction, disjunction, and negation. For example, given the two goals "run the marathon" and "do not get injured," the algebra offers the means of applying a disjunction to these two goal representations directly. Distributional reinforcement learning (Bellemare et al., 2017, 2023) offers another look at representing value, suggesting that decision-makers might benefit by representing the full distribution of future discounted reward. Despite this

17568765, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tops.70082 by Mark Ho - NYU Medical Central Library , Wiley Online Library on [27/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

representation capturing the same effective preference relation (in the sense that taking the expectation yields traditional value), it can still prove useful in generalization (Dabney et al., 2018). We speculate that composition, generalization, and learning efficiency are some of the plausible characteristics that might be desirable of goal representations. For a further examination of goal representations and their trade-offs, we refer to Davidson and Gureckis (2024).

2.3. Semantics preserving syntactic operations

In language, two syntactically distinct utterances (e.g., "Rex chased the cat" vs. "the cat was chased by Rex") can share the same or similar semantics (i.e., a situation where a dog named Rex is chasing a cat). How might this idea extend to the analysis of goal representations or task representations more generally? Suppose we adopted the account reviewed in the preceding section that formalized goal semantics as *preference orderings over policies* specified by a *value function*. Not only does this clarify what any single goal representation means or represents, it provides a basis for assessing the similarity of different goal representations.

For an example of a semantics-preserving operation for goal representations, consider *potential-based reward shaping* (Ng, Harada, & Russell, 1999). Given a long-term goal, reward shaping can be viewed as setting intermediate goals that provide short-term motivation. For instance, suppose one has the long-term goal of running the Boston marathon under 4 h, but to stay motivated they set the intermediate goal of improving their run time from the day before. A classic result from Ng et al. (1999) shows that shaping rewards must have a particular functional form (specifically, a difference of potentials); otherwise, the optimal behavior under the shaped rewards may differ from that under the original rewards. In our running example, this means that it is not enough to want to improve one's running speed over the previous day; one also needs to be averse to backsliding from the previous day. Crucially, without the additional goal to avoid backsliding, the optimal behavior is to repeatedly improve one's time and backslide, a behavior sometimes described as *exploiting a positive reward cycle* (Ho et al., 2019). From the perspective of our proposed syntax–semantics distinction, potential-based shaping is a syntactic operation (i.e., a constrained transformation of the original goal representation) that preserves semantic meaning (i.e., the optimal behavior).

Beyond goal representations themselves, work in artificial intelligence and cognitive science also explores how transformations of representations that interact with goals (e.g., world models) can be constrained to preserve optimal behavior. In particular, Grimm et al. (2020) introduce the notion of *value equivalent* representations to explain how deep model-based reinforcement learning systems learn world models that allow effective planning even if they are inaccurate. A similar principle underlies the *value-guided construal* framework, a normative account of how humans form ad hoc world models that trade-off value and complexity (Ho et al., 2022, 2023). To illustrate the idea of world models that induce similar goal semantics, imagine driving to a new destination in a familiar town. Even if you know all the bike paths, you are unlikely to include them in a world model for driving since those details will not change your route (cars cannot be driven on bike paths). Put another way, in the context of driving a car, the simpler model without bike paths preserves the value relations of the more complex model that includes bike paths. From the viewpoint

17568765, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tops.70082 by Mark Ho - NYU Medical Central Library , Wiley Online Library on [27/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

of a syntax–semantics distinction, these value relations are a semantic property, whereas representational complexity is a syntactic property.

Beyond value equivalence and value-guided construal for world models, we propose that drawing a clear analytical distinction between form and meaning can usefully frame research into goal representations. It is thus useful to ask, given some target semantic meaning (e.g., the value function defined by an extrinsic goal), what kinds of syntactic formulations preserve those semantics (i.e., subjective goal representations that induce roughly the same policy ordering). For example, one may have the goal of eating out less, but a semantically similar goal is to cook at home more. The choice of which of these two goal representations may matter: The latter may be a better starting point for planning since it is easier to start thinking about a single dish to cook rather than the many places one could avoid eating out at. More broadly, syntactically distinct but semantically equivalent representations often have different computational affordances. For example, as discussed, Markov rewards enable efficient dynamic programming (Puterman, 2014). Along similar lines, object-based representations may support more efficient planning while maintaining guarantees about value semantics (Olieslagers & Ho, in preparation).

3. Discussion

We have here drawn an analogy between the syntax–semantics divide in studies of language and logic to the way we conceive of goals. Using the standard goal formalisms of reinforcement learning (Sutton & Barto, 2018), this translates to (1) reward as goal atoms, (2) values as compositions of these atoms, and (3) preferences over policies as the representational content referred to by both rewards and values. To move between these three structures, we require additional bedrock commitments that amount to our axioms for goals. As in language and logic, we consider the syntactic operations that determine how we relate our atoms to sentences—that is, how reward is composed into value. Notably, this composition is often thought to occur both across *time* (in that value incorporates immediate and future reward) and *possibility* (in that value incorporates the likelihood of future rewards). To relate value to its representational content of a preference relation, we require further commitments that decide how precisely to form this relation in the form of so-called semantic axioms. In some cases, this is natural (and arguably so natural that it is taken as implicit): if we embrace vNM (von Neumann & Morgenstern, 1953), then we arrive at an equivalence between value as expected reward and preferences on outcomes. If we further want to impose structure onto our syntax, such as embracing Markovian rewards, then we require additional axioms to ensure this isomorphism is retained (Bowling et al., 2023). We saw how several recent accounts in the literature navigated these commitments in order to communicate new results about what kinds of content rewards of different forms can represent. While these accounts center around common formalisms of reinforcement learning, we again speculate that any account of goals that relates to experience will ultimately be expressible in terms of a preference relation of a certain kind. The questions are then: how is that goal represented by an agent (syntax), and what kind of preference does that goal representation refer to (semantics)?

17568765, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tops.70082 by Mark Ho - NYU Medical Central Library , Wiley Online Library on [27/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1
A nonexhaustive compendium of goal languages from the literature

Goal language	Syntactic affordances	Semantic expressivity
Markov reward (Sutton & Barto, 2018)	DP	Markov
Average reward (Mahadevan, 1996)	DP	Markov+
Risk-sensitive (Howard & Matheson, 1972)	DP	Markov+
Distributional (Bellemare, Dabney, & Munos, 2017; Bellemare, Dabney, & Rowland, 2023)	DP, Gen	Markov
Constrained (Altman, 2021)	DP	Markov+
Lexicographic (Shakerinava et al., 2025)	DP-	Markov+
Satisficing (Simon, 1956)	Unknown	Markov+
Blackwell optimal (Blackwell, 1962)	DP	Markov+
First-order logic (Ghallab et al., 1998)	DP, Prop	Markov
Temporal logic (Icarte et al., 2022)	DP, Prop	Markov+, Time, RegEx
Reward programs (Davidson et al., 2025)	Diverse	Markov+, Time, Comp

Note. Note on goal languages: Some are complete, some are partial). (DP = supports dynamic programming; DP– = does not support dynamic programming; Gen = facilitates generalization; Prop = propositional; Markov = expressivity of Markov reward; Markov+ = greater expressivity than Markov reward)

Open question: Axioms for goals? However, these are just the choices of relations between goal syntax and semantics the community has explored so far. We suggest that this framing highlights a significant conceptual opportunity to explore alternative compositional and semantic accounts of goals. Moreover, it positions us to scrutinize the needed relationship—if any—across the syntactic and semantic properties of a goal language. If we embrace a particular kind of compositional account of goal, what must this imply, if anything, about our semantic view of goals? We suggest that exploring the above questions provides a road-map to reaching bedrock clarity on the nature of goals. As summarized by Table 1, we suspect there is a profile of trade-offs in both the syntactic space (compositional affordances, for instance) and the semantic space (expressivity, for instance) that is worth elucidating.

Open question: The starting point as syntax or semantics? When conceiving of a goal, the proposal to think of goals in a syntactic or semantic form means that we have a choice: do we begin by committing to a goal as a *syntactic* or a *semantic* object, first? And, how does our theory and analysis need to change depending on our choice of this starting point? Using the language of rewards and preferences, this amounts to asking what affordances and problems we face if we select the representation as our starting point for thinking about goals, as opposed to the meaning. For example, we might start from rewards, which has arguably been the standard implicit view within reinforcement learning since the dawn of the field. When we start with reward, it offers convenient computational affordances. Specifically, a Markov reward function can be manipulated easily—through approaches like dynamic programming, we find simple mechanisms to compose these rewards into value (subject to suitable syntactic axioms). In fact, the premise that we can easily compute value from reward is so standard that we arguably have chosen our syntactic axioms *in the service of* making this computation straightforward.

17568765, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tops.70082 by Mark Ho - NYU Medical Central Library , Wiley Online Library on [27/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

As an alternative, we could have chosen others from Table 1, such as average reward. Yet, such quantities are known for being more difficult to compute than their discounted counterparts. In this way, we frame our syntactic commitments not just in terms of the intrinsic nature of what we can represent, but in terms of the computational properties offered by the syntactic structures we commit to (Davidson & Gureckis, 2024). If we were to instead start from preferences, we might find that questions about computational complexity or description length enter into our account of how to conceive of preferences in the right way (which has been somewhat absent thus far). We suggest these represent rich directions for further research.

Open question: Syntactic affordances and semantic expressivity as constraints on cognition? Lastly, a syntax–semantics distinction for goals can provide a useful analytical frame for investigating goal representations in human and animal cognition. In our view, this is not unlike studying the same cognitive process from, say, the computational versus algorithmic levels of analysis (Anderson, 1990; Marr, 1982). Specifically, certain properties of goals may be better understood in terms of their syntactic constraints (e.g., how well they afford efficient computation), while others in terms of their semantic constraints (e.g., how well they express rational behaviors in a certain environment). Characterizing how these distinct kinds of constraints interact with one another may also shed light on rational accounts of human cognition (Griffiths et al., 2024) and their computational plausibility (Rich et al., 2020).

3.1. Conclusion

Here, we explored the connection between several lines of research on goal representations and the syntax–semantics distinction from linguistics and logic. We propose that delineating the syntactic and semantic properties of goal representations can provide analytical clarity and guide future investigations of the goals that intelligent agents may pursue.

Acknowledgments

We would like to thank Logan Walls, Brandon Kaplowitz, as well as the reviewers for their valuable feedback on this work.

Note

1 To be clear, these two properties are not intended to be taken as necessary and sufficient conditions of goals, but rather, in our view, motivate the design and analysis of most formal accounts of goals.

References

Abel, D., Barreto, A., Roy, B. V., Precup, D., van Hasselt, H., & Singh, S. (2023). A definition of continual reinforcement learning. In *Advances in Neural Information Processing Systems*.

17568765, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tops.70082 by Mark Ho - NYU Medical Central Library , Wiley Online Library on [27/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Abel, D., Dabney, W., Harutyunyan, A., Ho, M. K., Littman, M. L., Precup, D., & Singh, S. (2021). On the expressivity of Markov reward. In *Advances in Neural Information Processing Systems*.

Altman, E. (2021). *Constrained Markov decision processes*. Routledge.

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Aumann, R. J. (1962). Utility theory without the completeness axiom. *Econometrica: Journal of the Econometric Society*, *30*(3), 445–462.

Ávila Pires, B., Rowland, M., Borsa, D., Guo, Z. D., Khetarpal, K., Barreto, A., Abel, D., Munos, R., & Dabney, W. (2025). Optimizing return distributions with distributional dynamic programming. *Journal of Machine Learning Research*, *26*(185), 1–90.

Baier, J. A., & McIlraith, S. A. (2006). Planning with first-order temporally extended goals using heuristic search. In P*roceedings of the AAAI Conference on Artificial Intelligence* (pp. 788–795).

Bakermans, J. J., Warren, J., Whittington, J. C., & Behrens, T. E. (2025). Constructing future behavior in the hippocampal formation through composition and replay. *Nature Neuroscience*, *28*(5), 1061–1072.

Ball, P. (2023). *Organisms as agents of evolution*. West Conshohocken, PA: John Templeton Foundation.

Bellemare, M. G., Dabney, W., & Munos, R. (2017). A distributional perspective on reinforcement learning. In *Proceedings of the International Conference on Machine Learning*.

Bellemare, M. G., Dabney, W., & Rowland, M. (2023). *Distributional reinforcement learning*. MIT Press.

Bellman, R. (1957). A Markovian decision process. *Journal of Mathematics and Mechanics*, *6*(5), 679–684.

Berridge, K. C., & Robinson, T. E. (2003). Parsing reward. *Trends in Neurosciences*, *26*(9), 507–513.

Blackwell, D. (1962). Discrete dynamic programming. *Annals of Mathematical Statistics*, *33*(2), 719–726.

Bowling, M., Martin, J. D., Abel, D., & Dabney, W. (2023). Settling the reward hypothesis. In *Proceedings of the International Conference on Machine Learning*.

Bratman, M. (1987). *Intention, plans, and practical reason*. Harvard University Press.

Chang, R. (2015). Value incomparability and incommensurability. In I. Hirose & J. Olson (Eds.), *The Oxford handbook of value theory* (pp. 205–224). Oxford University Press.

Cohen, J. D. (2017). Cognitive control: Core constructs and current considerations. In *The Wiley handbook of cognitive control*.

Dabney, W., Ostrovski, G., Silver, D., & Munos, R. (2018). Implicit quantile networks for distributional reinforcement learning. In Proceedings of the International Conference on Machine Learning.

Davidson, G., & Gureckis, T. (2024). Toward complex and structured goals in reinforcement learning. In *Finding the Frame: An RLC Workshop for Examining Conceptual Frameworks*.

Davidson, G., Todd, G., Togelius, J., Gureckis, T. M., & Lake, B. M. (2025). Goals as reward-producing programs. *Nature Machine Intelligence*, *7*(2), 205–220.

Dayan, P., & Niv, Y. (2008). Reinforcement learning: The good, the bad and the ugly. *Current Opinion in Neurobiology*, *18*(2), 185–196.

Decker, J. J. A. (2006). *Modality, rationalism, and conditionals*. PhD thesis, Massachusetts Institute of Technology.

Dennett, D. C. (1989). *The intentional stance*. MIT Press.

Dickinson, A. (1985). Actions and habits: The development of behavioural autonomy. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *308*(1135), 67–78.

Dickinson, A., & Balleine, B. W. (2000). Causal cognition and goal-directed action. In C. Heyes & L. Huber (Eds.), *The evolution of cognition* (pp. 185–204). MIT Press.

Dong, S., Van Roy, B., & Zhou, Z. (2022). Simple agent, complex environment: Efficient reinforcement learning with agent states. *Journal of Machine Learning Research*, *23*(255), 1–54.

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*(1–2), 3–71.

Gábor, Z., Kalmár, Z., & Szepesvári, C. (1998). Multi-criteria reinforcement learning. In *Proceedings of the International Conference on Machine Learning*, volume 98 (pp. 197–205).

Ghallab, M., Howe, A., Knoblock, C., McDermott, D., Ram, A., Veloso, M., Weld, D., & Wilkins, D. (1998). PDDL — The planning domain definition language, version 1.2. Technical Report CVC TR-98-003/DCS TR-

17568765, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tops.70082 by Mark Ho - NYU Medical Central Library , Wiley Online Library on [27/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

1165, Yale Center for Computational Vision and Control. *Prepared for the AIPS-98 Planning Competition Committee*.

Golinkoff, R. M., & Hirsh-Pasek, K. (2008). How toddlers begin to learn verbs. *Trends in Cognitive Sciences*, *12*(10), 397–403.

Griffiths, T. L., Callaway, F., Chang, M. B., Grant, E., Krueger, P. M., & Lieder, F. (2019). Doing more with less: Meta-reasoning and meta-learning in humans and machines. *Current Opinion in Behavioral Sciences*, *29*, 24–30.

Griffiths, T. L., Chater, N., & Tenenbaum, J. B. (2024). *Bayesian models of cognition: Reverse engineering the mind*. Cambridge, MA: MIT Press.

Grimm, C., Barreto, A., Singh, S., & Silver, D. (2020). The value equivalence principle for model-based reinforcement learning. *Advances in Neural Information Processing Systems*, *33*, 5541–5552.

Hammond, L., Abate, A., Gutierrez, J., & Wooldridge, M. (2021). Multi-agent reinforcement learning with temporal logic specifications. In *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems*.

Ho, M. K., Abel, D., Correa, C. G., Littman, M. L., Cohen, J. D., & Griffiths, T. L. (2022). People construct simplified mental representations to plan. *Nature*, *606*(7912), 129–136.

Ho, M. K., Cohen, J. D., & Griffiths, T. L. (2023). Rational simplification and rigidity in human planning. *Psychological Science*, *34*(11), 1281–1292.

Ho, M. K., Cushman, F., Littman, M. L., & Austerweil, J. L. (2019). People teach with rewards and punishments as communication, not reinforcements. *Journal of Experimental Psychology: General*, *148*(3), 520.

Howard, R. A., & Matheson, J. E. (1972). Risk-sensitive Markov decision processes. *Management Science*, *18*(7), 356–369.

Hutter, M. (2000). A theory of universal artificial intelligence based on algorithmic complexity. *arXiv preprint cs/0004001*.

Icarte, R. T., Klassen, T. Q., Valenzano, R., & McIlraith, S. A. (2022). Reward machines: Exploiting reward function structure in reinforcement learning. *Journal of Artificial Intelligence Research*, *73*, 173–208.

Jacobson, P. I. (2014). *Compositional semantics: An introduction to the syntax/semantics interface*. Oxford University Press.

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in Cognitive Sciences*, *20*(8), 589–604.

Johanson, L. (2003). Evidentiality in Turkic. *Studies in Evidentiality*, *273*, 290.

Kaelbling, L. P., Littman, M. L., & Cassandra, A. R. (1998). Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, *101*(1–2), 99–134.

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87–102.

Kumar, S., Marklund, H., Rao, A., Zhu, Y., Jeon, H. J., Liu, Y., & Van Roy, B. (2023). Continual learning as computationally constrained reinforcement learning. *arXiv preprint arXiv:2307.04345*.

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences*, *40*, e253.

Latham, G. P., & Locke, E. A. (1991). Self-regulation through goal setting. *Organizational Behavior and Human Decision Processes*, *50*(2), 212–247.

Lehman, J., Meyerson, E., El-Gaaly, T., Stanley, K. O., & Ziyaee, T. (2025). Evolution and the Knightian blindspot of machine learning. *arXiv preprint arXiv:2501.13075*.

Lewis, C. (1895). What the tortoise said to achilles. *Mind*, *4*(14), 278–280.

Littman, M. L., Topcu, U., Fu, J., Isbell, C., Wen, M., & MacGlashan, J. (2017). Environment-independent task specifications via GLTL. *arXiv preprint arXiv:1704.04341*.

Mahadevan, S. (1996). Average reward reinforcement learning: Foundations, algorithms, and empirical results. *Machine Learning*, *22*(1), 159–195.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco, CA: W. H. Freeman.

17568765, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tops.70082 by Mark Ho - NYU Medical Central Library , Wiley Online Library on [27/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Mattar, M. G., & Lengyel, M. (2022). Planning in the brain. *Neuron*, *110*(6), 914–934.
Mihatsch, O., & Neuneier, R. (2002). Risk-sensitive reinforcement learning. *Machine Learning*, *49*, 267–290.
Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*(1), 167–202.
Montague, R. (1970). English as a formal language. In B. V. Visentini (Ed.), *Linguaggi Nella Societá e Nella Tecnica* (pp. 188–221). Milan: Edizioni di Communita.
Newell, A., Shaw, J. C., & Simon, H. A. (1959). Report on a general problem solving program. In *IFIP Congress*, volume 256 (pp. 1959). Pittsburgh, PA.
Ng, A. Y., Harada, D., & Russell, S. (1999). Policy invariance under reward transformations: Theory and application to reward shaping. In Proceedings of the International Conference on Machine Learning.
Olieslagers, J., & Ho, M. K. (in preparation). Interactions between tree search and task simplification in planning.
Pierce, B. C. (2002). *Types and programming languages*. MIT Press.
Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526.
Priest, G. (1983). The logical paradoxes and the law of excluded middle. *Philosophical Quarterly*, *33*(131), 160–165.
Puterman, M. L. (2014). *Markov decision processes: Discrete stochastic dynamic programming*. John Wiley & Sons.
Rich, P., Blokpoel, M., de Haan, R., & van Rooij, I. (2020). How intractability spans the cognitive and evolutionary levels of explanation. *Topics in Cognitive Science*, *12*(4), 1382–1402.
Richens, J., Abel, D., Bellot, A., & Everitt, T. (2025). General agents contain world models. In *Proceedings of the International Conference on Machine Learning*.
Ringstrom, T. J. (2023). *Reward is not necessary: Foundations for compositional non-stationary non-Markovian hierarchical planning and intrinsically motivated autonomous agents*. PhD thesis, University of Minnesota.
Russell, S. J., & Norvig, P. (2010). *Artificial intelligence: A modern approach*. London.
Shakerinava, M., Ravanbakhsh, S., & Oberman, A. (2025). Beyond scalar rewards: An axiomatic framework for lexicographic MDPs. In *Advances in Neural Information Processing Systems (NeurIPS)*. Spotlight.
Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a rational and mechanistic account of mental effort. *Annual Review of Neuroscience*, *40*, 99–124.
Siewierska, A. (1984). *The passive: A contrastive linguistic analysis*. Croom Helm.
Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*(2), 129.
Sutton, R. S. (2004). The reward hypothesis.
Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT Press.
Tasse, G. N. (2020). A task algebra for agents in reinforcement learning (pp. 18–27). Master's thesis, University of the Witwatersrand, Johannesburg (South Africa).
Tasse, G. N., James, S., & Rosman, B. (2020). A Boolean task algebra for reinforcement learning. In *Advances in Neural Information Processing Systems*.
Velez-Ginorio, J., Siegel, M., Tenenbaum, J. B., & Jara-Ettinger, J. (2017). Interpreting actions by attributing compositional desires. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 39.
von Neumann, J., & Morgenstern, O. (1953). *Theory of games and economic behavior*. Princeton University Press.
White, M. (2017). Unifying task specification in reinforcement learning. In *Proceedings of the International Conference on Machine Learning*.
Wood, W., & Rünger, D. (2016). Psychology of habit. *Annual Review of Psychology*, *67*(1), 289–314.
Zadeh, L. A. (1988). Fuzzy logic. *Computer*, *21*(4), 83–93.

17568765, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tops.70082 by Mark Ho - NYU Medical Central Library , Wiley Online Library on [27/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License