Received: 29 June 2021 | Revised: 19 August 2021 | Accepted: 25 August 2021

DOI: 10.1111/mila.12405

WILEY

SUBMITTED ARTICLE

Inter-temporal rationality without temporal representation

Simon A. B. Brown

Department of Philosophy, Johns Hopkins University, Baltimore, Maryland, USA

Correspondence
Simon A. B. Brown, Department of Philosophy, Johns Hopkins University, 281 Gilman Hall, 3400 N. Charles Street, Baltimore, MD 21218, USA.
Email: simonabbrown@gmail.com

Recent influential accounts of temporal representation—the use of mental representations with explicit temporal contents, such as before and after relations and durations—sharply distinguish representation from mere sensitivity. A common, important picture of inter-temporal rationality is that it consists in maximizing total expected discounted utility across time. By analyzing reinforcement learning algorithms, this article shows that, given such notions of temporal representation and inter-temporal rationality, it would be possible for an agent to achieve inter-temporal rationality without temporal representation. It then explores potential upshots of this result for theorizing about rationality and representation.

KEYWORDS
diachronic rationality, inter-temporal choice, inter-temporal rationality, representation, temporal cognition, temporal representation

1 | INTRODUCTION

Inter-temporal rationality requires taking actions now which will not be immediately rewarding, but will result in larger benefits later. For example, it might involve forgoing food now to store a tool that can be used to obtain better quality food later, a feat Kabadayi and Osvath (2017) show in ravens. More generally, it involves consistently balancing one's welfare across different moments of one's lifetime. It might appear obvious that this requires a rich framework for understanding time, featuring concepts like IMMEDIATELY, LATER, and MOMENTS. I will show that it requires no such thing. In fact, a creature could be inter-temporally rational without having any representations with explicit temporal contents—such as attributing before and after relations or durations, or locating events at times—at all. Or, as I shall put it for convenience, inter-temporal rationality can be achieved without temporal representation.

This has implications for several debates. One is the nature of temporal representation—and representation generally. I remain neutral on exactly how to define representation. But my argument follows prominent recent discussions, such as Hoerl and McCormack (2019a) and Peacocke (2017), in assuming that temporal representation is demanding enough to contrast with mere temporal sensitivity. Such views face the challenge of articulating what *does* require representation. This challenge is sharpened if they imply that even inter-temporal rationality is possible without temporal representation. And if a demanding notion of representation cannot overcome this challenge in the case of time, this may cause problems for such views generally, including for Burge's (2010) influential approach.

Inter-temporal rationality — as I understand it here — without temporal representation also raises questions about the nature of rationality. I will be assuming that inter-temporal rationality is fundamentally about maximizing total utility over time. Some may wish to use this article's argument as part of a reductio of the picture of rationality it presupposes. To do so, one would need to specify exactly what is missing in the kind of decision-making I describe, and show that it does require temporal representation. Such a project would be highly informative.

This article also bears on issues surrounding the minds of different species: It shows that an animal could produce rational behavior in an interesting sense—including ravens' impressive delay gratification abilities demonstrated by Kabadayi and Osvath (2017)—without temporal representation, *given the learning mechanisms many animals have*. This has implications for the scope and significance of the traditional views that nonhuman animals lack rationality, and that they are "stuck in time," both of which have been thought to have profound implications for ethics and for our understanding of the mind and evolution.[1]

The basic argument of this article is that *reinforcement learning* (RL) algorithms can achieve inter-temporal rationality without relying on temporal representation. RL algorithms are an extremely important tool both in AI and as models of animal (including human) decision-making: It is likely that a great deal of behavior is due to brains implementing such algorithms. I will focus on two RL algorithms. The main argument (Sections 2 and 3) will show that reliably coming to act in a way that maximizes expected discounted utility over time is possible without representing time, using *temporal difference* (TD) learning. Section 4 will show that an even more sophisticated kind of inter-temporal rationality, reliably *immediately* acting in a way that maximizes expected discounted utility over time, is achievable sans temporal representation using model-based RL. Section 5 discusses objections and replies, and Section 6 explores some of the upshots of the result.

2 | THE MAIN ARGUMENT

My argument has the following overall structure (Section 3 will explain what the premises mean and defend them):

(P1) An agent is inter-temporally rational if they reliably act in ways which will, over time, maximize total discounted expected utility.
(P2) An agent that implements the RL algorithm TD will thereby reliably act in a way that will, over time, maximize total discounted expected utility.

[1]See Adamson and Edwards (2018) for the history of the rationality claim, and Roberts (2002) and Hoerl (2008) for the history of the "stuck in time" claim.

(P3) Implementing TD requires temporal representation only if representing one of the variables explicitly represented in TD—specifically the value function $V(.)$, the policy $\pi(.)$, or the current and next state S and S' (Section 3.2 explains what these are in detail)—requires temporal representation.
(P4) Having representational states with fixed temporally relevant functional roles does not require temporal representation.
(P5) Representing S, S', and $\pi(.)$ only requires having states with fixed temporally relevant functional roles.
(P6) Representing a variable that would be ultimately analyzed in partly temporal terms does not require temporal representation, provided the system simply treats the variable as unanalyzed.
(P7) $V(.)$ is a variable that would be ultimately analyzed in partly temporal terms, but TD can simply treat $V(.)$ as unanalyzed.

Therefore, (C) An agent can be inter-temporally rational without temporal representation.

(P4) and (P5) imply that representing S, S', and $\pi(.)$ does not require temporal representation; (P6) and (P7) imply that representing $V(.)$ does not require temporal representation either. So representation of S, S', $\pi(.)$, and $V(.)$ does not require temporal representation. But then, by P(3), TD can work without representation of time, which in turn implies by P(2) that an agent can act in a way that reliably maximizes total discounted utility without representation of time. Given P(1), this is enough for inter-temporal rationality without temporal representation.

I will now justify, and, where necessary, explain the different premises in turn.

3 | PREMISES

3.1 | An agent is inter-temporally rational if they reliably act in ways which will, over time, maximize total discounted expected utility

An important view of inter-temporal rationality, especially in formal models (including rational choice models in economics), is that an agent is inter-temporally rational insofar as they choose actions which maximize total discounted expected utility. One way of expressing this formally, if we call utility at the moment t "reward," symbolized R_t, and assume that future utility is discounted by a factor $\gamma \in [0,1]$ for every period in the future $t = 1$, $t = 2$, up to $t = n$, is to say that rationality requires agents choose actions $a_1, \ldots, a_n$ to maximize:

$$E\left[\sum_{t=1}^{t=n} \gamma^{t\text{-}1} R_t | a_1, \ldots, a_n\right] \tag{1}$$

Philosophers have developed alternatives to this view (Section 6.2), but it remains dominant. And with good reason. It elegantly captures the idea that rationality sometimes requires delaying gratification, but puts this idea in its proper place. It specifies a precise way of *trading off* present and future rewards, specifying how large a future reward (or extended stream of rewards) must be to justify foregoing some more immediate reward. And the idea that much of the time we are implicitly maximizing (Equation 1) can explain a great deal of behavior.

One way of calculating (Equation 1) would involve representing entire sequences of actions and predicted rewards, including explicitly representing *when* they all are predicted to occur, in order to discount rewards accordingly. By contrast, RL provides ways of ensuring actions reliably maximize (Equation 1), without such explicit temporal representation.

3.2 | An agent that implements the RL algorithm TD will thereby reliably act in a way that will, over time, maximize total discounted expected utility

RL is widely agreed to be an important part (although not the entirety—Hayden & Niv, 2021) of decision-making. Treating animals as performing RL algorithms can explain many aspects of both behavior and neural activity, and has correctly predicted previously unnoticed features of both, as summarized by Ludvig et al. (2011), Petter et al. (2018), and Sutton and Barto (2018, chapters 14 and 15). For example, it helped predict: how subtle differences in stimuli timing can reverse well-known classical conditioning effects like blocking (Kehoe et al., 1987); that dopamine activity corresponds to *unexpected* reward rather than reward simpliciter (Montague et al., 1996), and that this activity can be optogenetically manipulated to cause an animal to learn as if it were receiving unexpected reward and following an RL algorithm (Steinberg et al., 2013).

The general idea of RL algorithms is for a system to learn a policy—a function specifying which action to take for each possible state of the environment—which maximizes (Equation 1). To learn such a policy, the subject repeatedly takes different actions, moving into different states, and each time it does so, incrementally updates an estimate of the *value* of that state on the basis of the observed outcome on that occasion. The way this updating works ensures that, given enough iterations, these value representations track the expected contribution to total discounted rewards that one would get from moving into this state. This in turn allows the subject to improve their *policy* so that it better maximizes (Equation 1). It will be useful to see in detail how this works in a particular algorithm.

TD learning algorithms are one empirically and theoretically important variety of RL. Consider pseudo-code for one TD algorithm, Algorithm 1: see next page ("$x \leftarrow y$" means that the previous value of x gets replaced with y at that stage in processing):

The important parts for our purposes are these: TD arrives at an estimate for $V(S)$ (the value assigned to state S, on the assumption that after moving to S, the subject will take the actions dictated by $\pi(.)$) by repeatedly taking actions from S, and updating its estimate for $V(S)$ on the basis of the results—and likewise for every state. The value assigned to a state comes to track the extent to which taking that action or moving into that state will lead to immediate rewards, *in addition to* the extent to which it will put the subject in a position to take further rewarding actions in the future. So, when moving into a state leads to immediate rewards, TD will learn that this action is valuable; but an action might also be valuable due to its longer-run effects, despite low initial returns.

Although the true value which $V(.)$ is designed to track is the discounted sum of rewards from a sequence of actions, TD does not rely on representing sequences. Instead, the key to its tracking discounted future rewards lies in the specific way it performs incremental updates, given by line 9. The rationale for line 9 rests on two facts. First, the true value of S just is the expected immediate reward from S, plus the expected discounted value of the future stream of rewards from the next period onwards. Second, S' is the state in the next period, so the expected discounted value of the future stream of rewards from the next period just is what $V(S')$ is designed to track. line 9 therefore uses $V(S')$ (discounted because it will be reached 1 period after S) as a key part of an error signal—the expression inside the square brackets—which

ALGORITHM 1 Tabular TD for estimating $V(.)$ given π

1. Input: the policy π to be evaluated
2. Algorithm parameter: step size $\alpha \in (0,1]$
3. Initialize $V(s)$, for all $s \in \mathcal{S}^+$, arbitrarily except that $V(\text{terminal}) = 0$
4. Loop for each episode:
    5. Initialize S
    6. Loop for each episode:
        7. $A \leftarrow$ action given by π for S
        8. Take action A, observe R, S'
        9. $V(S) \leftarrow V(S) + \alpha\ [\ R + \gamma\ V(S') - V(S)\ ]$
        10. $S \leftarrow S'$
    11. until S is terminal. (Sutton & Barto, 2018, p. 120).

allows it to incrementally adjust $V(S)$ in the right direction. When $R + \gamma\ V(S') > V(S)$, that is, when the experienced reward and estimated value of the next state outstrip the current estimate of $V(S)$, the estimated value of S is increased proportionally (and decreased when the opposite holds). As more states are visited on more occasions, TD will converge on the correct values for $V(.)$—specifically, the value of moving into S and then following policy $\pi(.)$.

Upon improving this estimate, the policy itself can be improved, to one which always chooses to move into the state with the highest value of $V(.)$ available. As $V(.)$ was estimated on the basis of the old policy, it will now need to be relearned for the new policy. This process can be iterated, with policies and associated value estimates gradually improving until they converge on the optimum.[2] This convergence to an optimal policy is well known and studied in computer science—there are many proofs of such convergence for different RL algorithms, and a great deal of work is devoted to studying different algorithms' speed of convergence under different conditions.[3]

To illustrate, we can consider how TD would produce the future-directed behavior demonstrated in ravens by Kabadayi and Osvath (2017). Ravens learn that a tool can be used to obtain food from an apparatus. They subsequently choose the tool when it can be used *in the future* to open the apparatus, even when the apparatus is in another room. Furthermore, they learn that a specific human gives them food in exchange for tokens, and subsequently choose that token for future use, even before the human arrives. Ravens do this for delays between choice and payoff of 15 minutes (Experiment 1) or 17 hours (Experiment 2), even choosing the tool/token over immediate but lower-value rewards (Experiment 3).

TD can produce such behavior (see Lind, 2018, for detailed simulations of Kabadayi & Osvath's results with a related algorithm). The key reason for this is that ravens were given training with tools and tokens before they were tested on choosing them prospectively. TD might start out

[2]This is a simplification: For example, it may be required, especially in early stages of this iterative process, not to simply switch fully to the apparently value-maximizing policy, and instead to explore states that would not be visited under that policy. Complications of this sort do not require temporal representation, so I leave them aside.

[3]Sutton and Barto (2018) discuss convergence results throughout. For this algorithm specifically, see Sutton and Barto (2018, pp. 126-129, 139).

assigning value to different states arbitrarily. But during training, it would gradually increase the value of states progressively earlier in the chain leading up to the reward, and decrease the value of other states. First, it would increase the value of states immediately preceding reward—where the animal is using the tool on the apparatus or giving the token to the human—as visits to these states would result in a high value of R in line 9. It would then start increasing the value of states preceding *those* states—states where the subject was in possession of the tool/token and *near* to the apparatus/human—as visits to these states would be succeeded by a highly valued state, hence a high value of $V(S')$ in line 9. Ultimately, value would spread through the entire chain of states leading up to the reward—that is, to all states where the raven has a tool/token. States which never led to high rewards or highly valued states (i.e., states where the animal does not have a tool or token), meanwhile, would gradually decrease in value. On every visit to such a state S, its previously assigned value, $V(S)$, would be larger than $R + \gamma V(S')$, the sum of S''s experienced reward on that visit and the discounted value assigned to the state following S, until $V(S)$ decreased enough to reflect S's true (relatively low) value. Once the algorithm assigned higher values to states where the subject has the tool/token than to other states, it would reliably choose to move into those states—it would reliably choose the token/tool, even over small levels of immediate reward (R), and even when not in the immediate presence of the apparatus/human.

3.3 | Implementing TD requires representation of time only if representing $V(.)$, $\pi(.)$, S, or S' requires temporal representation

This premise can be split into two claims: (1) if some behavior or capacity is to motivate positing a representation of X, then the explanation of that behavior or capacity should involve computations which operate on or output representations with X as part of their contents; (2) if TD involves operating on or outputting representations with temporal content, then that temporal content must show up in its representations of $V(.)$, $\pi(.)$, S, or S'.

The first of these claims may sound trivial. However, many neuroscientists, and some philosophers, use "representation" in a deflationary way which does not respect this principle—often when giving an informal gloss on the workings of a system, but also in some theories of representation (see Burge, 2010, for detailed critical discussion of many such theories). For example, we could say "TD learning represents the whole stream of rewards spread over time rather than just immediate rewards" merely as a way of capturing the fact that a subject using TD will be *sensitive* to the whole stream of rewards spread over time — that is, that their behavior will vary appropriately with variations in total rewards.

This article concerns weightier uses of "representation," following recent prominent discussions of temporal representation. Peacocke (2017, p. 214) explicitly endorses the claim, which he draws from Burge's (2010) discussion of representation in other domains, that "[f]or an organism to be sensitive to a magnitude of a given type ... is not yet for the organism to represent that magnitude as such,"[4] and goes on to argue that "attribution of temporal representational content is correspondingly well-founded only if there are actions not fully explained by temporal features of proximal states ... but are explained by states of the creature whose content

[4]While Burge (2010) offers the most extended and influential recent discussion of the representation/sensitivity distinction in general, Burge (2010, pp. 518-529) claims that *temporal* sensitivities being "harnessed" for representation of other properties is sufficient for genuine representation (for which he is criticized by Peacocke, 2017 and by Gross, 2017).

involves duration and the past" (Peacocke, 2017, p. 218). Hoerl and McCormack (2019a, p. 4) adopt a similar notion, citing Peacocke.[5]

My argument does not rely on a specific account of the temporal representation/sensitivity distinction (like Peacocke's (2017) account in terms of "Representational Preservation"), or of representation in general (like Neander's (2017) or Shea's (2018) sophisticated teleosemantic accounts). The argument goes through on any account that endorses the claim that if some behavior or capacity is to motivate positing a representation of *X*, then the explanation of that behavior or capacity should involve computations which operate on or output representations with X as part of their contents, plus (P4) and (P6).

As for the claim that $V(.)$, $\pi(.)$, S, and S' are the only explicit representations in the above algorithm which are strong candidates for temporal representation, there are only two other potential candidates: the other represented variables (R, A, α, and γ); and mathematical functions of the variables, such as the entirety of the right-hand side of line 9. Any plausible arguments for any of these having temporal contents are stronger as arguments for $V(.)$, $\pi(.)$, S, and S'. To take them in turn:

R and A represent a reward and an action, respectively. There do not seem to be any reasons to think that these must be represented temporally, except reasons which also apply to the instantiation of states. α and γ, meanwhile, are just fixed parameters. α determines how much the RL algorithm updates its estimate of $V(S)$ in response to a single sample, and γ determines how much it discounts future events. One might think that α and γ have temporal content, in the sense that they change how the system behaves over time. But they do not track or vary with anything temporal: They do not vary at all. Neither does the algorithm calculate them or use them in different ways on the basis of their temporal aspect, instead using them in a fixed, automatic fashion. At best, they have the same status as $V(.)$: When we as theorists explain what they are, we have to advert to temporal features of the system's behavior, but the system itself does not do so.

As for mathematical functions of the represented atomic variables, we need not rule out *all* complicated mathematical functions of such variables counting as embedding temporal information: There does not seem to be any reason to include the specific functions involved in this algorithm as containing any more temporal content than their atomic components.

3.4 | Having representational states with fixed temporally relevant functional roles does not require temporal representation

(P4) and (P6) draw on commitments Hoerl and McCormack (2019a, 2019b) and Peacocke (2017) incur while arguing that various cases of temporally appropriate behavior do not imply temporal representation. They often proceed by showing that the behavior can be explained in terms of states with fixed functional roles that are partly delineated in terms of specific dynamics. I will explain what this means through examples.

Take coordinating perception, memory, and anticipation. We might think that this requires each of these processes to use contents with tense or markers like *past*, *present*, or *future*. Otherwise, we might think, subjects would frequently confuse representations pertaining to different times, respond now to long-past threats, and treat future plans as having already been carried out. Hoerl and McCormack argue otherwise. They argue that a "Temporal Updating System"

[5]McCormack & Hoerl (2017) point out that the view that infants are sensitive to duration without genuinely representing it dates back at least to Piaget.

could integrate information received in the past with information currently being perceived and with goal states, to produce appropriate actions, despite lacking temporal contents.

A temporal updating system keeps a model of the world, which retains information from past experience. It operates by "changing representations, rather than representing change" (Hoerl & McCormack, 2019a, p. 2). That is, when it gets new information contradicting its current model, it never concludes *the world has changed*: It simply changes its model and discards its previous information. If the model has it that there is a tree in a certain location because the subject visited that location in the past, but new information shows there is no tree there, the system does not conclude *there used to be a tree but it must have fallen down*; instead it just stops representing the tree as at that location. Its representations can be interpreted as *tenseless* (Hoerl & McCormack, 2019b, p. 52).[6]

The underlying move here is to build the temporal competence into the functional roles of the states involved, obviating the need for the states' *contents* to mark out the times they relate to. Hoerl and McCormack offer explanations of the same form for numerous more sophisticated cases. I will not discuss all of these, but it will be useful to get a flavor of how they deal with more complexity by considering sensitivity to duration.

Sensitivity to duration is extremely widespread. Many classic conditioning experiments involved rats and pigeons learning to respond to stimuli of particular durations, or to produce stimuli for certain durations (Gallistel, 1990, pp. 294, 301-306). And such sensitivity is often ecologically relevant. Peacocke (2017, p. 216) focuses on the Hawai'i 'amakihi returning to nectar sources after the optimal duration, which varies based on factors like how quickly the source refills and how likely others are to get there first, a behavior documented in detail in Gill's (1988) classic study of hummingbirds. Hoerl and McCormack focus on Clayton and Dickinson's (1998) finding that scrub jays return to sites where they have cached food only when that particular cache has not been left long enough to rot (given its contents: seeds, crickets, and worms decay at different rates), a behavior also documented in magpies (Zinkivskay et al., 2009), and chickadees (Feeney et al., 2009).

One way to return to the nectar source at the right time is to represent temporal content: to use a representation of the duration along with rate of refilling to estimate the current nectar level. But Peacocke (2017, p. 216) argues that there is an alternative: use a state representing nectar levels, which is governed by a mechanism that *automatically* changes its represented level at the right rate without any computation, much as a mechanism could automatically increase the level of water in a bucket at a constant rate without any computation. Similarly, to avoid searching at the sites of long-decayed caches, one need not rely on representing duration: McCormack (2001), Hoerl (2008), and Hoerl and McCormack (2019a, 2019b) argue that this is possible with a state that represents only the cache's contents and location, and, as part of its functional role qua worm-cache-representation, is governed by a timer so that it gets forgotten after the appropriate amount of time.[7]

This is not yet enough to explain *learning* to respond to durations. But Hoerl and McCormack (2019b, p. 52) argue that instead of having completely fixed dynamics, the timers governing such representations can be entrained—modified automatically to synchronize with

[6]Ismael (2017, p. 26-27) describes a similar system for different theoretical purposes.
[7]There may be more reason to attribute temporal representation to scrub jays than Hoerl and McCormack allow, as there is additional complexity in the behavior that they do not discuss (Clayton et al., 2003; Correia et al., 2007), and reason to suspect yet further complexity that has not been scientifically documented. However, the main point here is just illustration of what *could* be achieved without appeal to temporal representation in principle.

the environment—without representing duration. Suppose that when visits to the flower find it already completely full, the rate increases, and when the flower is not full, the rate decreases. Such a system will converge on the dynamics appropriate to that flower. Yet it would still be the case that the timing mechanisms "simply govern the updating and maintenance of elements of the creature's model of its present environment" (Hoerl & McCormack, 2019a, p. 14) rather than generating temporal representations that are parts of that model itself.

3.5 | Representing S, S', and $\pi(.)$ only requires having states with fixed temporally relevant functional roles

If this treatment is good enough for duration sensitivity, it is good enough for S, S', and $\pi(.)$. To see why, it will be useful to understand why one might think these involve temporal representation in the first place.

When explaining how the algorithm works, it is useful to gloss S and S' as "the previous state" and "the current state." Likewise, $\pi(.)$ is a function specifying which action to take *next*, given the *current* state. As such, we might think that the algorithm represents possible states *as* previous, current, and next.

However, we should distinguish between a description of a state that expresses its content, and a mere theorist's gloss. This distinction shows up in the cases in Section 3.4: We *could* say that the 'amakihi's representation is of the nectar being available "now and not a moment ago". But for Peacocke, it does not include this temporal content explicitly, and such talk is merely a way of talking about the dynamics of the nectar-representation that are part of its functional role. The description of S, S', and $\pi(.)$ as representing the previous/next/current states only needs to be a gloss of this kind. Their functional roles can be fixed with respect to temporal properties, with S and S' *always* representing whatever the current state and the previous states happen to be, and $\pi(.)$ always representing a function from states to actions, which is always consulted in the same way, for the same purposes. We can imagine the whole system being governed by a regular oscillator (like how Hoerl and McCormack and Peacocke think birds' representations might be governed by a timer), such that every cycle, the represented state of the environment in the S' slot is copied into the S slot, overwriting the previous entry there, before perception automatically determines which state of the environment should be put in the S' slot, after which the algorithm is looped through again.

Of course, one *could* implement TD using representations with temporal content rather relying on fixed functional roles; but one *need not* do anything like this to achieve inter-temporal rationality: fixed functional roles are good enough.

3.6 | Representing a variable that would be ultimately analyzed in partly temporal terms does not require temporal representation, provided the system simply treats the variable as unanalyzed

A theorist's gloss might misconstrue a state's content not just through picking up on dynamic aspects of the state's functional role, but through representing what I will call a "covertly temporal entity": an entity which should ultimately be analyzed by philosophers and scientists partly in temporal terms, but which is not represented as such by the subject. As in Section 3.4, we can see this through examples.

A covertly temporal property for human perception might be pitch. Ultimately, pitch is frequency of sound waves *per second*. It is not a pure temporal property: It matters that it concerns *sound* waves. But its temporal component is essential too. Perceptual experience, however, does not separate these components out. The fact that pitch has this temporal component is a surprise to most physics students.

As it is for pitch, so it *could* be for all rates. It is crucial to Hoerl and McCormack's project of arguing that temporal representation is unique to adult humans that they have an account of both sensitivity to and representation of rates which does not invoke temporal representation. Sensitivity to rates is extremely important and widespread in the animal kingdom (Gallistel, 1990, chapter 11). Indeed, Gallistel and King (2010, pp. 226-240) argue that all classical conditioning involves rate representation, and Petter et al. (2018) make related claims about RL, which we will discuss in Section 5. Van Duijn et al. (2006), Lyon (2015), and Bechtel and Bich (2021), and others who argue that even bacteria engage in cognition, emphasize microbes' complex forms of sensitivity to rates, like *Escherichia coli*'s sensitivity to the rate of change of chemical concentrations controlling their movement. So Hoerl and McCormack need to either say that many of these cases of rate sensitivity do not involve rate representation, or that rate representation does not involve temporal representation, even though rates involve time. For our purposes, instances of their making the latter move are more important.

They make it when replying to Pan and Carruthers's (2019) claim that optimal foraging theory implies that animals represent time, because it implies their representing rates of reward.[8] Hoerl and McCormack say this:

> [T]hink of the dial on a car's speedometer. This acts as a representation of a rate, a rate related to the locations the car occupies at different times, but it does not act as a representation of a succession of events happening at different times. Arguably, people also use it in ways that don't need to involve any reasoning about other times, such as when they simply look at the speedometer, notice that they are breaking the speed limit, and take their foot off the gas. In a similar way, we want to suggest that animals may be sensitive to the rate of reward at different locations, but in a way that does not involve any reasoning about different moments in time. (Hoerl & McCormack, 2019b, p. 60)

I think there are two moves here. One is the emphasis on representation of *times* as opposed to *temporal properties*; this move is not available to my argument, as my target includes all pure temporal properties. The other is the point that a speedometer is an example showing one can represent a rate without representing it as temporal—without breaking it down into some variable that changes over time (such as miles) and a unit of time (such as hours).

One might agree that, in principle, one could *have* a representation of a rate without temporal representation, yet question how a system could *reliably produce* such representations accurately without representing time as such. Surely sensitivity to rates of change requires either differentiation of a function of time, or dividing the difference between a start and end value of a variable over a temporal interval by that interval's duration? The answer is no. Fixed dynamic functional roles for states with nontemporal contents can do the job. Suppose that the variable

[8]They are not the only ones to make this point, for example, "[a]ll sorts of memories of actions, sensations, or distances could become embodied representations, which are intrinsically related to time, without having a 'pure' representation of the dimension itself" (Osvath & Kabadayi, 2019, p. 41)

whose rate of change we want to measure is temperature. Have a component that generates readings of temperature, governed by an oscillator so that it does so every second.[9] Each reading generates a representation m. Suppose also that each period, just before it is replaced by a new reading, m is copied and a new state is formed with the most recent reading. Then $n(m - n)$/(period of the oscillator) gives the average rate of change during the period. Because the oscillator has a fixed period, $m - n$ itself will track the average rate of change.

We can bolster the point that rates do not require temporal representation with two further cases. Firstly, it is plausible that ordinary objects should *ultimately* be analyzed partly in terms of their persistence conditions: What it is to be a teapot is partly to maintain a certain form over time in such a way as to hold tea under normal conditions. And representing a teapot as a teapot arguably requires some kind of implicit sensitivity to this fact. But such sensitivity could be very implicit indeed; and as such, it would be misleading to say that a representation of something as a teapot is a temporal representation. Secondly, we can consider a nontemporal case: We might think that the number seven is ultimately to be defined as $1 + 1 + 1 + 1 + 1 + 1 + 1$; but this need not be appreciated (except, perhaps, very implicitly) by the subject that represents there being seven biscuits in the tin. It would be misleading to say that such a representation is of (or even *partly of*) the successor relation or +.

3.7 | *V*(.) is a variable that would be ultimately analyzed in partly temporal terms, but TD can simply treat *V*(.) as unanalyzed

If this treatment is good enough for rate sensitivity, it is good enough for V(.). V(.) is a variable that would be ultimately analyzed in partly temporal terms. If we as theorists were to analyze what V(.) represents, we might say something like: "the value of different states—the expected discounted rewards over time that would be achieved from entering each possible state." However, just like for our experience of pitch or the speedometer's representation of speed, this is just our gloss, relying on distinctions the subject does not (and potentially cannot) make, and V(.) itself is not explicitly temporal. We might think of the content of $V(S)$ as something more like an unarticulated [EXPECTED-VALUE-OVER-TIME-GIVEN-π-OF S] than [EXPECTED VALUE OVER TIME GIVEN π OF S]. A subject can run an RL algorithm without being capable of appreciating the analysis of V(.) or in any way decomposing it into its temporal components beyond the "analysis" implicit in performing the computations involved in TD learning—which I have already argued do not involve temporal representation. Just as it would be misleading to say that the speedometer represents time, that ordinary object representations have temporal contents or that an ordinary representation of seven represents the successor relation, it is misleading to say that V(.) has temporal content. As for the question of how it can be calculated without temporal representation, the TD algorithm provides an answer.

It might be thought that V(.) is not simply atomic for such a system—there is some analysis of its components, in the sense that it is broken down into $V(S) + \alpha\,[\mathrm{R} + \gamma\,\mathrm{V}(S') - V(S)]$ in line 9. However, any "temporal" component here is not a representation reusable in other contexts as temporal—it is in the relationship between S' and S that does not extend beyond this single computation.

[9]Peacocke (2017, p. 217) argues in more detail (along basically the same lines as for sensitivity to duration) that sensitivity to phases in regular cycles can be achieved without temporal representation.

Indeed, Hoerl and McCormack (2019b, p. 53) do allude to the idea that value representations could implicitly encode temporal information when replying to concerns Osvath and Kabadayi (2019) raise, suggesting that they agree that value representations are not temporal, although they do not go into detail.

4 | IMMEDIATE RATIONALITY

This completes the basic argument that inter-temporal rationality can be achieved without temporal representation. This is a surprising and interesting result in its own right, but it does have various limitations, deriving from limitations of TD. Some of these relate to the fact that to learn $V(.)$ for every state, the system needs to attempt the task repeatedly. During these attempts, it will decidedly not act inter-temporally rationally. It will take actions which would seem irrational to us, because it needs to actually explore states whose pitfalls we could foresee without visiting them. Suppose the task being learned is playing chess: *We* might not need to ever try out intentionally putting our queen in danger with no discernible benefit, to know that this would be a bad move; but the TD algorithm would require try it out—and then do it again for good measure. Algorithm 1 will only approach an optimal policy in the long run, at least for large state-spaces.

However, this particular problem is not endemic to RL: there are various ways of supplementing the algorithm which can help overcome it, without introducing temporal representation. One strategy is building in generalization of various kinds, so that updates to the value of one state automatically inform the algorithm's valuation of relevantly similar states. Another, which I will focus on here, is to use a *model* to simulate the environment. Given a simulation of what would happen if the subject visited a certain state, RL can determine the optimal policy (given its model of the world) before it has taken any real actions at all. Not only can it achieve this even more impressive kind of inter-temporal rationality than the TD algorithm in Section 3, but model-based RL is an important topic in AI and computational neuroscience, so it will be worth going through why it, too, does not require temporal representation.

A model is something which outputs representations of S' and R given representations of S and A. We can use the very same TD algorithm as in Section 3, only using simulated states, actions and rewards (i.e., line 8 will change from "Take action A, observe R, S'" to "Simulate taking action A, observe predicted R, S'"). If the model matches the actual world (if it has the correct probabilities of outputting different values of S' and R), then repeating this process enough times will lead $V(.)$ to converge on the correct values, for the very same reasons as repeating actual interactions with the world.

Does this require temporal representation? Most of the process for learning $V(.)$, given the model, is the same as in Section 3, so there are only a few potential reasons why temporal representation might have been introduced: using the model, representing the model, or learning the model.

Using the model might be thought to involve temporal representation, because of the following contrast: above, I said that TD learning did not require temporal representation because it always involves representing the present and immediately previous states, and these can be taken care of by fixed functional roles. Models, by contrast, allow simulation of a whole temporal chain of states in order, stretching into the future. However, while some RL algorithms *do* use models to generate entire chains, which they represent as such and operate on in virtue of their temporal properties, this is not necessary: At any given stage of processing, the system need only represent only two states, a reward, and an action, and overall it just needs to explore all the different states, without needing to explore them in any particular order.

Similar points apply to representing the model: all it requires is a fixed function, with slots for S, S', R, and A, much like $\pi(.)$ has fixed slots for S and A. It need only represent two states (and an action and reward) at once, and it can do so with slots that have distinctive functional roles, rather than representing the relationship between the two states as temporal. These functional roles will correspond to the fact that we want the model to output states and rewards according to those that will temporally follow a given action and state, but this correspondence need not be represented.

Finally, one might worry that an accurate model of the world can only be *learned* using temporal representation. But, while writing out and discussing specific algorithms for learning models would take us too far afield here, it should be intuitive already that learning about which state-action pairs predict which state-reward pairs can be achieved through learning based on machinery already shown to not imply temporal representation—representations of pairs of states (or quadruples of two states, a reward and an action) in functionally defined slots, automatic calculation of and response to error signals, and so forth.

5 | OBJECTIONS AND REPLIES

RL is sometimes presented as involving temporal representation. Petter et al. (2018, p. 911) state flatly, "[t]he computational goal of RL is to maximize future rewards, and this depends crucially on a representation of time"; and Hayman and Huebner (2019) appeal to animals' doing RL as an example of their using temporal representation. Different ways of understanding such claims have different upshots for the argument of this article.

First, there are claims that all RL algorithms, including TD, rely on temporal representation, which appear to be based on glosses of the different components of RL, of precisely the kind I already argued above do not reflect genuine temporal representation, at least on the demanding notion of representation presupposed here. For example:

> Model-free systems compute forward-looking predictions, track discrepancies between experienced and predicted rewards, and adjust future predictions to accommodate such discrepancies. Model-based strategies store a model of the world that specifies when a sequence of actions is expected to yield reward and compute decisions on this basis...each kind of system relies on forward-looking expectations about which actions are likely to be rewarded, as well as retained representations of what has worked in the past. (Hayman & Huebner, 2019, p. 26)

Second, RL sometimes operates over state representations which include explicitly temporal information. States might be defined in terms of the duration of some stimulus, just as they can be defined by the presence of a particular shape or length. This is one way of equipping RL to learn solutions to timing-specific problems such as peak interval procedures, temporal bisection tasks, and tasks involving integrating temporal relationships between different pairs of events. Much of Petter et al. (2018) is devoted to the use of RL in such tasks. However, this does not pose a worry for the above argument: *If* such tasks require temporal representation, this is not because RL *in general* requires temporal representation, or because inter-temporal rationality requires temporal representation. Rather, it would reflect independent facts about these specific tasks.

A different worry in the vicinity is that *empirically plausible* models of inter-temporal decision-making in animals involve temporal representation. The specific form of TD learning discussed above is only one flavor of RL, and there are many others. Not all variations here pose any extra need for temporal representation: For example, using value functions that apply to state-action pairs rather than to states. But some have prima facie claims to involving temporal representation, and might be realistic models of some brain activity: For example, algorithms involving replaying entire sequences of actions. However, it is important to emphasize that the claim of this article is not that animals never use temporal representation, but that temporal representation is not *required* for inter-temporal rationality.

The most serious worry here is that focusing on actual brains might bring out ways in which *all* RL algorithms, including TD, in fact rely on temporal representation in ways that were tacitly suppressed by the above presentation. Realistic cases might be thought to show that either state representations or the update rule require temporal representation after all.

One argument for this would start by claiming that state representation relies on perception, and then combine this view with the claims that we mainly perceive continuous, dynamic entities like processes, trajectories, motion and the change of different variables (Grush, 2007; Ismael, 2017, pp. 25, 29), and/or that to perceive even static objects, we need to integrate continuously changing views of those objects (Hayman & Huebner, 2019 raise this as an issue for Hoerl & McCormack, (2019a, 2019b); for related discussion see Burge, 2010, p. 445). As it stands, this worry is not specific to inter-temporal rationality, and it relies on a view about perception which Hoerl and McCormack would deny. But it does bring up a more specific worry: that the version of TD I discussed above requires discretely segmented states, occurring at discretely segmented time periods. What if it turns out that any genuine mind, as opposed to simple computer, has to work in continuous time rather than such segmented states?

One reply to this worry is that there is evidence that we *do* automatically segment our stream of experience into events for certain purposes (Zacks & Tversky, 2001; Clewett et al., 2019).[10] But more importantly, there are versions of RL that use continuous rather than discrete changes in states. The key question will be whether *they* imply temporal representation. Thoroughly answering this would require detailed discussion of particular algorithms (which are considerably more mathematically involved than the TD algorithm above). But the short answer is that the main candidates for temporal representation in these algorithms are rate representations, and Section 3.6 already argued that these do not imply temporal representation.

A closely related worry is that all RL algorithms have some limitation or other. Convergence proofs rely on various assumptions. So it might be questioned whether they *really* achieve inter-temporal rationality, as opposed to *conditional rationality*, *rationality given that these assumptions hold*. Such assumptions are important, and in Section 6.2, I will articulate some of them and consider the prospects of taking them as the starting points for an account of behavior that does require temporal representation. For now, even without getting into the details of such conditions, we can see that the worry about whether RL counts as achieving rationality given its reliance on such conditions is misplaced. We can see that they are often met, at least approximately, from the fact that RL is used successfully in many real-world scenarios by brains and

[10]There is a further question about whether this event segmentation is possible without the perceptual system operating over representations of time, but this is independent of the issue of whether inter-temporal choice requires temporal representation, given that it has state representations.

artificial applications of machine learning. Furthermore, relying on background conditions to achieve rationality does not rule out actually achieving rationality when such conditions are met: Even procedures which no one would question as counting as implementing inter-temporal rationality rest on *some* conditions to successfully do so. For example, directly calculating (Equation 1) will only result in inter-temporal rationality in a creature capable of performing such calculations accurately, who has formed the expected values they are calculating with accurately, and who is calculating over a state and action space that carves up the world appropriately.

6 | UPSHOTS OF THE ARGUMENT

6.1 | Inter-temporal rationality?

Maximizing total expected discounted utility, while a prominent account of inter-temporal rationality, is not the only account on the market. If one thinks that the claim that inter-temporal rationality without temporal representation is implausible, then one might treat the argument here as a consideration in favor of an alternative account of rationality. Here I briefly sketch how the discussion might go for three alternative conceptions of inter-temporal rationality, in order to bring out my argument's potential consequences.

The most flatfooted response along these lines is that genuine rationality requires being *aware* that you are maximizing total expected discounted utility over time, and doing so *intentionally*. This view would effectively stipulate that inter-temporal rationality would require representation of time (as well as of maximization, rationality, etc.). One issue with this view is that it is extremely rare that ordinary humans think in anything like these terms explicitly, even if they have a better grasp of time than a TD algorithm does.

One might instead think that inter-temporal rationality requires that subjects decide the *shape* of utility's distribution across their lives. Hoerl and McCormack (2019a, 2019b, p. 15) suggest that the ability to do this might be one of the main benefits of temporal representation (albeit in the context of a discussion which arguably underestimates the sophistication of inter-temporal choice achievable through RL). Slote (1982) argues that it is at least rationally permissible to have preferences about the overall shape of utility, and Velleman (1991) links this to preferences about the narrative structure of one's life, while Schechtman (1996, 2011), Taylor (1989), and others have argued for connected claims about having a unified, autonomous self across time at all involving appreciating such narratives. However, these claims have been criticized (e.g., by Strawson, 2004); and it remains far from clear how necessary for rationality such ways of thinking about inter-temporal choice really are.

A related view has it that inter-temporal rationality requires other kinds of consistency over time, including certain kinds of stability in preferences and choices, and sticking to plans (Bratman, 2018). Having and resolutely sticking to a plan may require representing a sequence of actions as occurring in a temporal order, and representing my having made a plan in the past tense.

Irrespective of the outcome of these debates, all should agree that maximizing utility over time, even without realizing it, is *relevant* to, if not exhaustive of, inter-temporal rationality. On any plausible account of inter-temporal rationality, delaying gratification will be an important part of the story, and that is one of the things that can be done without temporal representation.

6.2 | Temporal representation?

My argument treats representation as a demanding notion, in line with Peacocke (2017), Hoerl and McCormack (2019a, 2019b), and the general attitude to representation in Burge (2010). This does not require a specific positive account of representation, but does require more than some prominent attempts at naturalizing representation, such as simple covariation theories and at least some versions of teleosemantics. For example, Montemayor (2013) argues on the basis of a broadly teleosemantic account, and Gallistel (1990) argues on a functioning isomorphism view, that temporal representation is implied by many of the cases that Hoerl and McCormack think merely show sensitivity (cf., Montemayor, 2019; Pan & Carruthers, 2019; Viera & Margolis, 2019).

Advocates of a deflationary account might treat the result that inter-temporal rationality is possible without temporal representation, *given a demanding account of representation*, as part of a reductio of demanding accounts. But advocates of deflationary accounts will face their own problem, of showing exactly why representation is implicated in TD, while maintaining "representation" as a useful notion. After all, nearly all organisms behave appropriately with respect to time. It is crucial to their functioning and often evolutionarily selected for that they undergo regular cycles which synchronize with one another and with the environment (Gallistel, 1990, pp. 221-222; Montemayor, 2013, p. 35). And as mentioned above, *E. coli*'s movement is sensitive to rates of change of chemical concentrations. Should all of this be counted as temporal representation? If so, what do we gain by describing it as such? If not, what precisely is the difference between bacteria going through a regular cycle and a brain doing so while running a TD algorithm? Many who are tempted by *somewhat* deflationary accounts of representation will be reluctant to go so far.

Even if deflationists cannot answer this challenge, demanding views of temporal representation still face a challenge which is sharpened by the finding here: Specifying what *would* be enough for temporal representation. What distinctive capacities can temporal representation underpin which we cannot explain through fixed functional roles and unanalyzed covertly temporal representations? If such a view implies that even inter-temporal rationality is possible without temporal representation, we might worry that it is likely to leave temporal representation implicated only in extremely specific sorts of temporal capacities, if any at all. And the argument above is suggestive that many other kinds of competence with respect to time could be achieved using fixed functional roles and representation of covertly temporal properties, given the mileage that can be got out of these in the context of RL. What other algorithms could incorporate such tricks, and to what ends?

Peacocke (2017) and Hoerl and McCormack (2019a, 2019b) do sketch positive accounts of what temporal representation can buy us. But both warrant further discussion, given how far we can push temporally competent behavior without (by their lights) temporal representation. Hoerl and McCormack (2019a, 2019b) emphasize the ability to represent orders of objective events which differ from the order of subjective experience of those events.[11] But it is not clear that this is a particularly significant achievement, given that one of the main uses they suggest for it is for the contentious form of choice concerned with the shape of utility over time discussed in Section 6.1. Furthermore, it may turn out to be possible to represent orders of

[11]Elsewhere they emphasize prima facie different capacities, like causal reasoning that respects the principle that causes precede effects (Hoerl, 2008), although Hoerl and McCormack (2011) closely link *this* to an account of representation of objective order closer to the (Hoerl & McCormack, 2019a, 2019b) account.

objective events which differ from the order of subjective experience via RL or other styles of algorithm building on the sorts of tricks in Section 3. Peacocke, meanwhile, points to "Representational Preservation": Representing the world as being a certain way an explicitly specified length of time ago, in addition to integrating all such representations into a representation of the total state of the world both past and present. One worry for this account is that it is unclear, without more detail, if there really are any behaviors which strictly require such a system, and could not be achieved by a sufficiently complex algorithm making use of the kind of tricks discussed in the context of RL.

In Section 5, I alluded to conditions under which RL does not reliably produce expected utility maximization. Rather than treating such conditions as showing RL *never* achieves inter-temporal rationality, a view I rejected in Section 5, such conditions could serve as the starting point for a demanding account of temporal representation. Perhaps temporal representation can be defined in terms of *achieving inter-temporal rationality even when these conditions are violated*.

Many such conditions relate to RL's relying on a well-chosen state space. The set of states which the system represents needs to instantiate an effective way of carving up the world, or RL can produce suboptimal results. There is a cluster of ways that state-spaces can be poorly chosen for RL. The most obviously time-related is violating the *Markov assumption*. This requires that the current state on its own is the best predictor for what state will come next—that including the history of states leading up to this current state would not add predictive power. When this assumption is violated, any update rule which, like line 9 in Section 3.2's TD algorithm, only takes into account a single time period, systematically fails to incorporate available information about states' values.

The suggestion, then, would be to define temporal representation in terms of being able to achieve inter-temporal rationality in the absence of the Markov property. Fully developing this approach would require exploring algorithms which achieve this, and carefully considering whether such algorithms require temporal representation. However, not only is such a project well beyond the scope of this article, but it might reasonably be suspected of resting its account of temporal representation on excessively technical details, which do not themselves directly connect to general accounts of representation or of time. This suspicion might motivate a very different approach, that reconsiders the aim of finding specific patterns of behavior that can only be explained in terms of temporal representation. Such an approach might look at more than just behavior, or rethink the notion of explanation at issue and its relationship to representation.

This article is not claiming that either austere or deflationary views of temporal representation are decisively ruled out; rather, I hope to have deepened our understanding of what is at stake between the two camps, by showing that an important part of inter-temporal rationality can be achieved without (demanding) temporal representation, and by showing some of the tricks that allow for this—tricks that can be generalized to other cases.

7 | CONCLUSION

I have shown that, if we assume an austere (though not unmotivated) account of temporal representation, and a utility-maximizing account of inter-temporal rationality, then inter-temporal rationality is achievable without temporal representation. RL algorithms can produce inter-temporally rational actions while relying only on representations which an austere account would not consider temporal. This seems to imply that at least one of the following three claims must be true: the austere account of temporal representation is too demanding; the utility-maximizing account of inter-temporal rationality is not demanding enough; or the relationship between inter-temporal rationality and temporal representation is very different to what many have assumed.

ACKNOWLEDGEMENTS

For helpful comments and discussion on earlier versions of this material, I would like to thank two anonymous reviewers, in addition to Nemira Gasiunas-Kopp; Matthew Heeney; Jenann Ismael; John Morrison; Christopher Peacocke; Ian Phillips; participants and audience members in the American Philosophical Association Pacific Division online meeting in 2021 and the *International Symposium Thinking in and about time: A dual systems perspective on temporal cognition* in Milan in 2019; and attendees of: the Perception and Mind Lab meeting at Johns Hopkins University; the NYU Philosophy of Mind Discussion Group; and PoPRocks workshop in New York.

ORCID

Simon A. B. Brown https://orcid.org/0000-0002-5606-6979

REFERENCES

Adamson, P., & Edwards, G. F. (Eds.). (2018). *Animals*. Oxford University Press.

Bechtel, W., & Bich, L. (2021). Grounding cognition: Heterarchical control mechanisms in biology. *Philosophical Transactions of the Royal Society B*, *376*, 20190751.

Bratman, M. (2018). *Planning, time, and self-governance*. Oxford University Press.

Burge, T. (2010). *Origins of objectivity*. Oxford University Press.

Clayton, N., & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature*, *395*(6699), 272–274.

Clayton, N. S., Yu, K. S., & Dickinson, A. (2003). Interacting cache memories: Evidence for flexible memory use by Western scrub-jays (*Aphelocoma californica*). *Journal of Experimental Psychology: Animal Behavior Processes*, *29*(1), 14–22.

Clewett, D., DuBrow, S., & Davachi, L. (2019). Transcending time in the brain: How event memories are constructed from experience. *Hippocampus*, *29*(3), 162–183.

Correia, S., Dickinson, A., & Clayton, N. S. (2007). Western scrub-jays anticipate future needs independently of their current motivational state. *Current Biology*, *17*(10), 856–861.

Feeney, M. C., Roberts, W. A., & Sherry, D. F. (2009). Memory for what, where, and when in the black-capped chickadee (*Poecile atricapillus*). *Animal Cognition*, *12*(6), 767–777.

Gallistel, C. R. (1990). *The organization of learning*. MIT Press.

Gallistel, C. R., & King, A. P. (2010). *Memory and the computational brain: Why cognitive science will transform neuroscience*. Wiley-Blackwell.

Gill, F. B. (1988). Trapline foraging by hermit hummingbirds: Competition for an undefended, renewable resource. *Ecology*, *69*(6), 1933–1942.

Gross, S. (2017). Perception and the origins of temporal representation. *Pacific Philosophical Quarterly*, *98*(S1), 275–292.

Grush, R. (2007). Time and experience. In T. Müller (Ed.), *Philosophie der Zeit* (pp. 27–44). Klosterman.

Hayden, B. & Niv, Y. (2021). The case against economic values in the orbitofrontal cortex (or anywhere else in the brain). *Behavioral Neuroscience*, *135*(2), 192–201.

Hayman, G., & Huebner, B. (2019). Temporal updating, behavioral learning, and the phenomenology of time-consciousness. *Behavioral and Brain Sciences*, *42*(e254), 25–26.

Hoerl, C. (2008). On being stuck in time. *Phenomenology and the Cognitive Sciences*, *7*, 485–500.

Hoerl, C., & McCormack, T. (2011). Time in cognitive development. In C. Callender (Ed.), *The Oxford handbook of philosophy of time* (pp. 439–459). Oxford University Press.

Hoerl, C., & McCormack, T. (2018). Animal minds in time: The question of episodic memory. In K. Andrews & J. Beck (Eds.), *The Routledge handbook of philosophy of animal minds* (pp. 56–64). Routledge.

Hoerl, C., & McCormack, T. (2019a). Thinking in and about time: A dual systems perspective on temporal cognition. *Behavioral and Brain Sciences*, *42*(e244), 1–16.

Hoerl, C., & McCormack, T. (2019b). Authors' response: Temporal updating, temporal reasoning, and the domain of time. *Behavioral and Brain Sciences*, *42*(e278), 51–69.

Ismael, J. (2017). Passage, flow, and the logic of temporal perspectives. In C. Bouton & P. Huneman (Eds.), *Time of nature and the nature of time* (pp. 23–38). Springer.

Kabadayi, C., & Osvath, M. (2017). Ravens parallel great apes in flexible planning for tool-use and bartering. *Science*, *357*(6347), 202–204.

Kehoe, E. J., Schreurs, B. G., & Graham, P. (1987). Temporal primacy overrides prior training in serial compound conditioning of the rabbit's nictitating membrane response. *Animal Learning & Behavior*, *15*(4), 455–464.

Lind, J. (2018). What can associative learning do for planning? *Royal Society Open Science*, *5*(11), 180778.

Ludvig, E. A., Bellemare, M. G., & Pearson, K. G. (2011). A primer on reinforcement learning in the brain: Psychological, computational, and neural perspectives. In E. Alonso & E. Mondragón (Eds.), *Computational neuroscience for advancing artificial intelligence: Models, methods and applications*. Hershey, PA.

Lyon, P. (2015). The cognitive cell: Bacterial behavior reconsidered. *Frontiers in Microbiology*, *6*, 264.

McCormack, T. (2001). Attributing episodic memory to animals and children. In C. Hoerl & T. McCormack (Eds.), *Time and memory: Issues in philosophy and psychology* (pp. 285–313). Oxford University Press.

McCormack, T., & Hoerl, C. (2017). The development of temporal concepts: Learning to locate events in time. *Timing & Time Perception*, *5*(3-4), 297–327.

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *The Journal of Neuroscience*, *16*(5), 1936–1947.

Montemayor, C. (2013). *Minding time: A philosophical and theoretical approach to the psychology of time*. Brill.

Montemayor, C. (2019). On the human uniqueness of the temporal reasoning system. *Behavioral and Brain Sciences*, *42*(e266), 38–40.

Neander, K. (2017). *A mark of the mental*. MIT Press.

Osvath, M., & Kabadayi, C. (2019). A theory stuck in evolutionary and historical time. *Behavioral and Brain Sciences*, *42*(e268), 40–41.

Pan, S., & Carruthers, P. (2019). No doing without time. *Behavioral and Brain Sciences*, *42*(e270), 42–43.

Peacocke, C. (2017). Temporal perception, magnitudes and phenomenal externalism. In I. Phillips (Ed.), *The Routledge handbook of philosophy of temporal experience* (pp. 213–224). Routledge.

Petter, E. A., Gershman, S. J., & Meck, W. H. (2018). Integrating models of interval timing and reinforcement learning. *Trends in Cognitive Sciences*, *22*(10), 911–922.

Roberts, W. (2002). Are animals stuck in time? *Psychological Bulletin*, *128*(3), 473–489.

Schechtman, M. (1996). *The constitution of selves*. Cornell University Press.

Schechtman, M. (2011). The narrative self. In S. Gallagher (Ed.), *The Oxford handbook of the self* (pp. 394–418). Oxford University Press.

Shea, N. (2018). *Representation in cognitive science*. Oxford University Press.

Slote, M. (1982). Goods and lives. *Pacific Philosophical Quarterly*, *63*(4), 311–326.

Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., & Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nature Neuroscience*, *16*(7), 966–973.

Strawson, G. (2004). Against narrativity. *Ratio*, *17*(4), 428–452.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.

Taylor, C. (1989). *Sources of the self*. Harvard University Press.

van Duijn, M., Keijzer, F., & Franken, D. (2006). Principles of minimal cognition: Casting cognition as sensorimotor coordination. *Adaptive Behavior*, *14*(2), 157–170.

Velleman, J. D. (1991). Well-being And time. *Pacific Philosophical Quarterly*, *72*(1), 48–77.

Viera, G., & Margolis, E. (2019). Animals are not cognitively stuck in time. *Behavioral and Brain Sciences*, *42*(e277), 50–51.

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*(1), 3–21.

Zinkivskay, A., Nazir, F., & Smulders, T. V. (2009). *What-where-when* memory in magpies (*Pica pica*). *Animal Cognition*, *12*(1), 119–125.

How to cite this article: Brown, S. A. B. (2022). Inter-temporal rationality without temporal representation. *Mind & Language*, 1–20. https://doi.org/10.1111/mila.12405