CellPress

Review

Deciding How To Decide: Self-Control and Meta-Decision Making

Y-Lan Boureau,[1] Peter Sokol-Hessner,[1] and Nathaniel D. Daw[1,*]

Many different situations related to self control involve competition between two routes to decisions: default and frugal versus more resource-intensive. Examples include habits versus deliberative decisions, fatigue versus cognitive effort, and Pavlovian versus instrumental decision making. We propose that these situations are linked by a strikingly similar core dilemma, pitting the opportunity costs of monopolizing shared resources such as executive functions for some time, against the possibility of obtaining a better outcome. We offer a unifying normative perspective on this underlying rational meta-optimization, review how this may tie together recent advances in many separate areas, and connect several independent models. Finally, we suggest that the crucial mechanisms and meta-decision variables may be shared across domains.

Trends

Meta-decisions are decisions about how decisions are made. Many recent models in different domains have conceptualized meta-decision dilemmas as pitting more carefully computed decisions against automatic defaults, including goal-directed versus habitual responses, deliberative versus heuristic choices, and controlled versus impulsive actions.

These recent models show that many puzzling decision patterns as well as phenomena of self-control and conflict can be understood as rational arbitration that balances the potentially better outcomes of more considered decisions against the higher costs of such consideration.

A central cost of deliberation across many seemingly separate domains has been proposed to be the opportunity cost of occupying shared resources over time. Common decision variables and mechanisms may guide these allocations across many such domains.

The Choice To Exercise Control

Smart people constantly fail to 'do the right thing'. We procrastinate, eat unhealthy food, and generally defeat our own goals. But why? Such behaviors are particularly vexing for influential normative and decision theoretic perspectives on cognition, which conceptualize decision making as maximizing long-term obtained reward. If we are optimizing, why should we ever be 'of two minds' about anything?

We advance here a unifying normative perspective on a range of situations involving conflict or self-control (very broadly construed), including automaticity, deliberation, and habits; Pavlovian reflexes; emotion regulation; fatigue and cognitive effort; and **learned helplessness** (see Glossary). The linking idea, versions of which have recently been proposed more or less separately in several of these sub-areas [1–5], is that true optimization requires 'meta-optimization' that accounts for the benefits and costs of the internal processes employed in making decisions. Putting aside context-specific details, the underlying decision architecture and trade-offs involved are strikingly similar, and in each case rely on balancing the benefits of higher rewards (from making a more optimal decision) against the increased costs of arriving at that choice [1–9]. Thus, the decision is not only over the possible outcomes but also over the nuts and bolts of the internal decision processes themselves (or, 'setting the switches' [10]). We suggest that some simple mechanisms and decision variables may be widely shared across these domains.

Adopting as an illustration the widespread view that the brain houses (at least) two distinct **decision controllers** (e.g., [4,11]), meta-optimization entails choices such as selecting the controller that performs the optimization and allocating to it resources such as time [12–14], then selecting the final outcome according to the preferences of that controller.

[1]New York University, 4 Washington Place, NY 10003, USA

*Correspondence: ndaw@princeton.edu (N.D. Daw).

 http://dx.doi.org/10.1016/j.tics.2015.08.013

Normative Meta-Decision Making

A typical laboratory model of self control asks how long you will keep your hand in unpleasantly cold icewater [15]. Doing so might ensure that the experimenter pays you, but requires you to constantly inhibit the urge to escape the noxious sensation by withdrawing your hand. The core dilemma of this and many other control tasks is, in effect, whether to simplify the choice process. Often, as here, this simplification consists of releasing an automatized, prepotent response (e.g., removing your hand from the icewater). Overriding such a default response to make any other, more contextually appropriate choice (e.g., not removing your hand so as to eventually get the payment) is what gives the situation a flavor of conflict and self-control.

In general, choosing among options involves computing and comparing their expected values. This computation can be laborious (such as when contemplating different sets of moves in chess), but may also be accomplished with shortcuts such as pre-wired defaults or frugal heuristics [16]. Heavier resource use may (or may not) yield a more rewarding ultimate outcome, but this comes at the cost of occupying resources that could otherwise have been used for something else, potentially with its own rewards. The net value of selecting any decision process is the value of the final outcome of that process, minus the cost of the resources used to reach a decision (e.g., evaluation time [4,5,17]), and/or implement it (e.g., executive attention for inhibiting a prepotent response [18]).

The crux of the matter is to figure out whether the benefits of a more laborious decision process are worth the corollary resource costs [5,12], which in turn itself requires in some way estimating those meta-decision variables using evidence from the current environmental context and reinforcement history. Accordingly, as would be expected from a rational economic decision, the amount of control that subjects choose to allocate in the laboratory is sensitive to variation in factors such as reward [19] and to manipulations that induce beliefs about whether control is a limited or unlimited resource [20]. In addition, recent findings have shown that cognitive control is deployed as would be predicted by normative models of the investment of a costly resource: (i) monetary incentives make people avoid cognitive 'work' less [7], (ii) allocation of time between a wage-earning demanding task and a non-wage-earning easy task responds to wage manipulations as predicted by labor supply theory [9], (iii) the subjective cost of cognitive effort can be measured in monetary units through repeated economic choice experiments [2], and (iv) activity in cognitive control-related brain areas reflects the integration of the costs and benefits of control [3]. We propose here that the logic of balancing costly resources against better outcomes can be extended beyond cognitive control to any context that involves the possibility of overriding a default response.

The Benefits of Control

Investing resources can lead to a better outcome in many contexts, for instance by inhibiting maladaptive fear [21] or inappropriate prepotent responses [18], thus allowing more accurate estimates of outcome value leading to ultimately more rewarding choices [22], or by keeping in mind contextual information that helps to make responses both faster and more accurate [19].

A common theme across these examples is that choosing the most rewarding option depends on accurately knowing the value of the candidates. Default actions such as reflexes or habits represent rough approximations that are insensitive to changes in context. Computational models of how resource use translates into gain (i) assume that a particular amount of resource use buys better information, for instance by considering a larger set of possible outcomes or attributes or by allowing a higher number of computational operations [4,16,23], (ii) specify mechanisms by which investing resources may achieve more accurate responses [24,25], (iii) describe how search in complex cognitive representations could secure a better outcome [22,26], or (iv) use information theory to determine how much control should be required to supply the information that justifies overriding the default response [27].

Glossary

Central executive: a flexible system regulating and coordinating cognitive processes.
Decision controller: a system that chooses actions, typically taking into account environmental circumstances (sensory inputs, state of the world, etc.).
Ego depletion: an experience-dependent impairment of performance in tasks requiring self-control or cognitive control, usually observed after having performed a preliminary demanding task.
Habitual controller: a decision controller that maps a context or a stimulus to a learned response after extensive training (e.g., turn right at the end of the block). A habitual controller is more flexible than a Pavlovian controller in that the learned response can be arbitrary instead of drawing from a limited repertoire of innate responses.
Goal-directed controller: a decision controller that evaluates available actions in terms of a prediction about their outcomes (e.g., turning right leads to the subway station). Such evaluations are more flexible than the choices of a habitual controller, but may require deliberation over multiple steps.
Learned helplessness: a change in behavior induced by exposure to situations in which an agent has no control over outcomes. The agent no longer attempts to cope with the environment (e.g., accepting electric shocks instead of escaping).
Marginal value theorem: a theorem describing optimal behavior in certain foraging problems. A forager sequentially visiting depleting resources (such as fruit trees) should optimally seek a new resource when the rate of return falls below the opportunity cost of time spent foraging there, which is given by the overall average reward rate in the environment.
Model-based control: a computational theory for goal-directed control in which options are evaluated using a learned model of the consequences of actions.
Model-free control: a decision controller that relies on an aggregated summary of past returns of actions, without using a model of the particular consequences of the actions. Model-free control is a

In this way, the dilemma of whether to override an automatic response is analogous to the familiar explore–exploit tradeoff [28] where subjects must figure out the value of different options by sampling them over trials (exploration), and balance this sampling against earning the most reward by choosing the options that seem best on the basis of current knowledge (exploitation). The benefit of exploration is the value of the information gained in terms of enabling more rewarding choices on future trials [29,30], and the limited resource is the set of available trials.

More quantitatively mapping the functional relationship between resource use and outcome value [31] is difficult: beyond the perennial problem of identifying the subjective reward of an outcome [32–35], it is hard to quantify the amount of resources that were mobilized to secure it. Current empirical strategies include metabolic measures such as expired gas analysis [36], behavioral economic procedures that gauge how much money subjects need to be paid to choose a demanding task over an easy one [2], and direct brain function measurements contrasting activity in regions of the prefrontal cortex believed to subserve executive functions during easy and hard decision tasks [6].

The Costs of Control

Viewing failure to exert control as irrational often focuses solely on the gain side of a more costly decision, while overlooking the cost [37].

Although cognitive resources do have intrinsic costs (e.g., the metabolic cost of firing spikes) – as the old economic adage goes, all costs are ultimately opportunity costs. In other words, the cost equates to what one could have obtained by spending the same resource some other way (Box 1). Thus the crucial computations of resource cost usually hinge on comparing the values that could be obtained from different possible uses. Using a resource is costly if this involves foregoing another beneficial use, and cheap if it does not (Figure 1).

The importance of opportunity costs has been recognized in several contexts such as foraging [38], free operant responding [39], temporal discounting [40,41], goal-directed versus habitual control [4], action sequence chunking [42], and cognitive effort [1]. In the former cases, such as foraging, the principle of lost opportunity is physical: one cannot eat from two bushes at once. Less obviously, cognitive resources that can only be used by one process at a time, such as attention or working memory, pose analogous time allocation problems. Research in all these domains generally pits one option against another within a single context (e.g., foraging in the same patch vs leaving to look for a better one). However, the limited resources at stake are

proposed computational theory of habitual control.

Pavlovian controller: a decision controller that evokes innate responses to stimuli that are predictive of biologically important consequences (such as salivating to a bell that predicts food).

Sample complexity: the number of examples an algorithm needs to obtain a satisfactory estimate of a quantity of interest (e.g., how many times does a customer need to go to a restaurant to provide a reliable rating?).

Box 1. Opportunity Costs and Average Reward

Opportunity costs arise when multiple mutually-exclusive actions are available at a given moment. For example, a child has $2 to purchase an ice cream cone. That money could purchase several different cones each of which cost $2, but purchasing one necessarily means not purchasing another. Thus the child should consider not only what they might purchase but what they might not. Calculating opportunity costs can be computationally daunting, especially if the option space is large. In our example, this might arise if the child considers not only ice cream cones but also other confections. To get around this challenge of exhaustively representing every possible alternative, for problems that involve the allocation of time one can simplify the computation by maintaining a running average of rewards received over time – the average reward rate. That average reward rate is an estimate of the opportunity cost of time [39]. If an action has a greater value than the average reward rate, then it is 'worth' taking (a decision rule formalized in the '**marginal value theorem**' [38]) whereas, conversely, actions which deliver worse-than average returns are not worth the time they take.

In some particular scenarios, such as animal foraging, this simple comparison rule is optimal; in many others it may be a reasonable approximation. In particular, this decision rule is myopic: it does not consider the possibility that investing resources now (e.g., training an underperforming employee) might carry enhanced returns later, and it considers investment as all-or-nothing for some period (measuring what one could earn overall during that period, and not the particular value of occupying e.g., some slot of a working memory store).

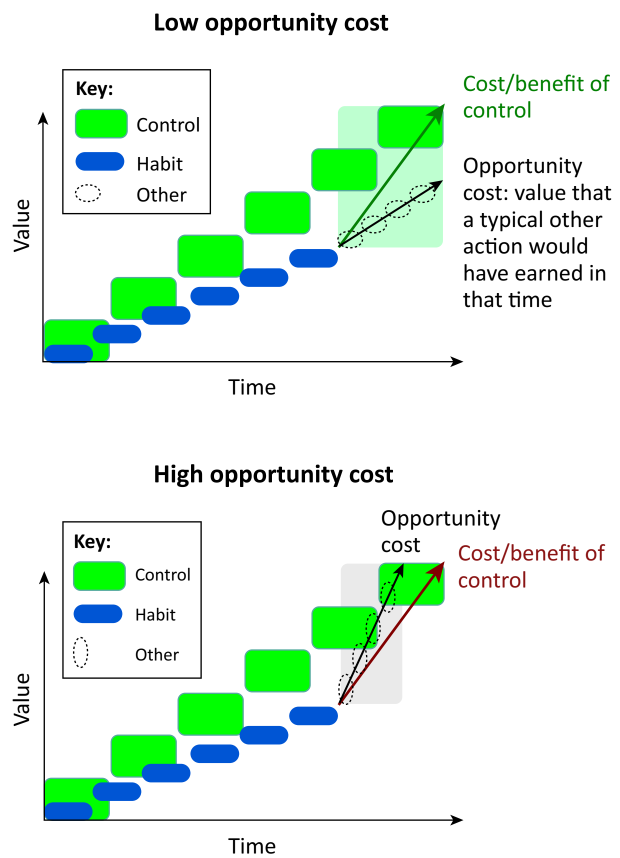

Figure 1. Opportunity Cost. The figure illustrates opportunity costs in a forced-choice task in six trials in which the agent can choose between a cheap, low-value habitual response, and a more-costly, more-valuable controlled response. Note that the six green boxes (Control) finish at a higher value but also require longer time than the six blue boxes (Habit). The question to the organism is which set of boxes (controllers for action) to choose – the green or the blue? To answer this question we must also know the opportunity cost of time. The opportunity cost captures what else could be done after the task is completed, and could be estimated as the long-term average reward in the environment (depicted here as an 'other' action). In the low opportunity cost case, control buys more additional value than the current average reward (green versus black arrow), and control is thus advantageous despite the additional time it takes. In the high opportunity cost case, the foregone other action would be more valuable than the additional value attained with control, and habit should therefore be preferred (such that the 6 actions are completed sooner, and the 'other' action can be performed instead).

shared across domains: executive function and even lower-level modules such as the visual system are ubiquitously useful. Thus, competition is not only between tasks directly offered by the experimenter but also with any latent options such as daydreaming, planning lunch, or worrying about how to make ends meet if money is scarce [43].

This type of reasoning implies that the costs of control ultimately arise from competitive allocation of cognitive resources that are shared across multiple possible uses, but can only be used for one at a time [1]. These shared and limited resources include executive resources such as working memory, attention, and the **central executive** [44,45] (Box 2). Dependence on these resources makes many processes vulnerable to contexts that affect their capacity. For instance, stress imposes a strain on executive functioning, and thereby causes emotion regulation to break down [46] and deliberative, model-based choices to weaken [47]. Tasks from many different areas involving executive function have been shown to cross-influence one another [48,49], and cross-predict performance [50,51].

Box 2. Limits of Executive Function

The existence of a shared, multi-purpose limited central executive has been hypothesized ever since early experimental work demonstrated limits on the number of simultaneous executive tasks that could be performed without interference between them [44,98]. Executive functions enable flexible adaptation to changes in the environment and motivation switching [99], support reactive, proactive, and counterfactual inferences [100], and their engagement has been suggested to correlate with the subjective feeling of mental effort [25]. The central idea is that there is a limit to the number of items that can be attended to, maintained, or cognitively manipulated at any given time, making executive function by definition a limited resource. Thus, the allocation of executive function to a given task almost always involves opportunity costs (Box 1).

Balancing the Costs and Benefits of Control

Figuring out the exact costs and benefits of control requires difficult computations and knowledge of all the contingencies of the task at hand. This raises a problem of infinite regress, particularly because such meta-decisions themselves are supposed to assess whether or not it is worthwhile to simplify decision computations.

One shortcut is to suppose that the brain might use rough quantities that approximately capture the costs and benefits of control. Two candidates for such approximate meta-decision variables seem broadly applicable: the average reward rate and controllability.

When devoting one's shared resources to a task for a period of time, the opportunity cost of that time is given in many settings either exactly or approximately by the long-term average reward rate [39]. This represents the amount of reward one would expect to earn, on average per unit of time, and thus may be a reasonable proxy or bound on the reward foregone by occupying reward-relevant resources for that time. When the world is richer, time is more costly. The average reward as opportunity cost has been proposed to govern decisions about the temporal allocation of resources in physical effort and vigor [39], prey selection [52], patch foraging [53], deliberation [4], action sequencing [42], and time discounting [40]. Thus the brain may track the average reward rate and apply it across many types of decisions, including meta-decisions. The logic of this coarse measure is to treat allocation as all-or-nothing – thereby equating the cost of occupying the resource over time to the overall cost of the time itself. For this reason, it does not apply directly to more granular resource allocation decisions, such as how much of some graded resource (e.g., working memory) is to be occupied, or whether to spend some resource that improves decisions while also speeding them up. Even in these cases, it may still provide a useful proxy provided that the contribution of individual resources to reward uniformly scales up as the overall reward increases.

The benefits of control, meanwhile, quantify how much more reward would be gained by optimizing more carefully. A rough proxy for this is given by various measures of the controllability of the environment [23]. There are several ways of quantifying controllability, all related to the extent to which one's actions determine one's outcomes. In this context, controllability can capture the advantage of a carefully considered action over a random one – for example, the difference between the reward for the best action versus a randomly chosen action (Figure 2). If this difference is small, it tends to be a waste to spend resources optimizing: one might as well just choose randomly or use a default response. Again, it has been argued that the brain tracks this quantity and uses it to weigh decisions, as in depression and learned helplessness [54], where repeated experience with uncontrollable situations leads to subsequent passivity in other tasks. Altogether, a cartoon summary of this shortcut would be:

$$\text{Net benefit of control} \approx \text{Controllability} - \text{Average reward}$$

More generally, we would expect the willingness to exert control or effort, across many domains, to increase with controllability but decrease with average reward. Of course, these rough proxies

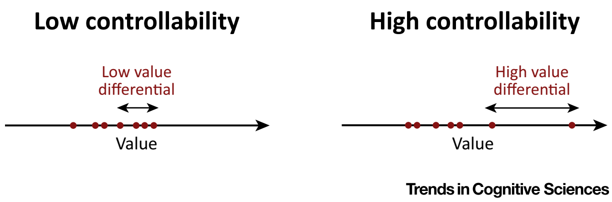

Figure 2. Controllability. A simple measure of controllability is the difference between the values of the best and average achievable outcomes. The figure illustrates the values associated with a sample of observed actions. Values are similar in the low controllability case, but, in the high controllability case, careful choice optimization can lead to a much better outcome because the best value is further removed from the typical case.

might be refined with decision variables more suited to particular situations or contexts, for which these more general quantities might serve as priors. For instance, it appears that uncertainty about the value of a particular action (over and above controllability more generally) affects meta-decisions in situations such as exploration [30] and habits [22]. Indeed, while we sketch here a strategy of identifying situations in which control is generally more advantageous or disadvantageous, other theoretical work in psychology envisages that the brain chooses among approximate computation strategies by learning the empirical performance of particular heuristics in particular situations [5,16]. These two approaches may well be complementary.

The general point is that meta-decision variables of this type represent broad features of the environment that an organism can easily estimate from experience (e.g., by averaging received rewards) and use to guide the allocation of control, generalizing across many situations or tasks. This suggests accounts for a family of phenomena, including **ego depletion** [1] and learned helplessness [54], in which experience with environmental statistics drives changes in later control allocation. Any change in the decision context affecting either the expected value of control or the expected cost of resources may change the decision process and the way its parameters are set. For instance, if two options differ in their resource requirement, a preference reversal may be caused by a change in the estimated resource opportunity cost even if the true opportunity cost has in fact stayed the same.

Examining the expected benefits and costs of resource investment sheds light on four related types of puzzles. First, why are there controllers and motivational states characterized by low resource use (e.g., the **Pavlovian controller** [55], fatigue [56,57], depression [41,58])? Second, what should guide arbitration between different decision styles (e.g., **model-based** versus **model-free control** [22,59–61], impulsive versus controlled [18])? Third, how should organisms trade-off resource versus performance within a given process (e.g., time and attention devoted to an executive task [9], time taken to make a choice [4])? Finally, how and why should decision strategies change depending on previous experience?

Three Example Domains of Meta-Decision Making

Goal-Directed versus Habitual Decision Making

Probably the best understood example of rational meta-decision making, empirically and theoretically, is for decisions about whether to deliberate. The expected value of an action can be retrieved from aggregate past experience or worked out more prospectively by enumerating the consequences of the action. These two strategies are termed model-free and model-based decision-making, and are closely related to the categories of habitual and goal-directed responding from behavioral psychology [22,59–61]. For example, in a typical laboratory situation, a rodent presses a lever for some food. If the animal has been overtrained, lever-pressing can become a habit; that is, a type of learned automatic response directly associated to context, which bypasses the association with the outcome of food and is therefore no longer sensitive to desire for food. Conversely, model-based deliberation relies on identifying outcomes and estimating their values. These two types of control can be distinguished by testing whether the animal will still press the lever when the food is unwanted, such as after an aversion has been induced by pairing it with illness.

We can view perseverative lever-pressing in this case as reflecting a meta-decision to engage the habit: in other words that the extra cost of extended model-based deliberation was not expected to be offset by sufficient benefit beyond that provided by the habitual response.

On the benefit side, evaluating an action in a model-based fashion is more accurate in terms of what computer scientists call **sample complexity**: it uses all available data in near-optimal fashion [22]. The most obvious cost of constructing model-based values is time, because values have to be estimated on the fly [62]. Model-based values could be obtained by searching a tree of the possible consequences of action sequences [22], where the order in which possibilities are examined may itself also be optimized according to the expected value [26]. Experimental evidence in humans suggests that deliberatively constructing value requires time and attention [63–65], and that many different quantities capturing outcome statistics are computed in different parts of the brain [66].

Model-free choices ('habits'), conversely, are fast but sloppy. They represent learned automatic responses which can simply be triggered without further computation. As in the case of sated rats working for food, this lack of deliberation can lead to errors.

Normative models of arbitration between **habitual** and **goal-directed controllers** have been proposed very much along the lines of our general sketch above. In particular, it has been suggested that the brain employs average reward rate as an estimate of the opportunity cost of time, and compares it to the expected benefit of more accurate value estimates to determine whether a goal-directed or habitual controller is optimal [4,23]. This model explains a great deal of data about the circumstances under which either type of control dominates, such as the emergence of habits with overtraining: once enough samples have been seen, model-free responses are expected to be well-calibrated and the reduced accuracy gap with model-based ones no longer (in expectation) justifies the extra deliberation costs.

Subjective Effort and Control

Although the work reviewed above concerns the time devoted to value estimation, an additional cost of controlled action not yet considered is the ongoing opportunity costs due to occupying executive resources needed to execute a planned course of action. This happens, for instance, if an action requires inhibiting a prepotent response in an ongoing manner, as with holding one's hand in icewater or engaging in a sustained way with onerous work [18,67].

Recognizing this has led to a model of arbitration between controlled or effortful versus default responses [1] that is otherwise formally fairly parallel to models of arbitration between goal-directed and habitual responses [4]. Another appeal of this model is that it may help to explain systematic changes over time in subjects' allocation of control (e.g., so-called 'ego depletion' [48]). In ego depletion experiments, subject performance on a control-demanding probe task is measured following some initial, usually different, 'depleting' task. Typically, performance on a variety of probes (from icewater endurance to anagram unscrambling) is reduced if the initial task itself demanded control, as though such tasks require some common (but unidentified) depleting resource. A rational meta-decision account may explain this phenomenon via experience-dependent learning about cost or benefit estimates, without appealing to an unidentified depleting resource, as in the more traditional strength model of self-control [68,69].

That said, compared to operant choice in rodents, in many human cognitive effort experiments the costs and benefits of controlled action are less objectively manipulated and harder to quantify. Because objective performance-dependent payments are often not given, understanding the benefits often comes down to assumptions about the demand characteristics of complying with the requests of the experimenter. The estimation of the opportunity cost of executive resources is also rather quantitatively unconstrained in this setting, as is how either of these quantities would interact with experience to give rise to the phenomenology of ego-depletion. In principle, ongoing learning about controllability, or average rewards, or both, might lead to ego depletion phenomena. The average reward rate as the opportunity cost of time is only directly applicable if the decision to engage is all-or-nothing, for example whether to perform a task versus daydream or quit. In tasks where the agent can decide how much to apply herself (e.g., focusing hard versus half-heartedly on solving puzzles), it might be possible to extend this idea to track the average reward rate obtained per unit of executive resource use. Quantities of this type may be related to theories of learned industriousness and self-efficacy [70–72].

Cognitive Emotion Regulation and the Pavlovian Controller

Another example of using executive resources to inhibit a prepotent response is cognitive emotion regulation [21,73,74]. On the benefit side, cognitive emotion regulation can make behavior more adaptive [75]. As for cost, it appears to depend on similar executive circuits as the other behaviors discussed here [51,75–77].

Inhibiting Pavlovian responses is a closely related example. Many examples of self-control in the more colloquial sense, such as inhibiting the consumption of unhealthy food or impulsivity in the 'marshmallow test', involve suppressing Pavlovian response tendencies such as that to approach and consume appetitive stimuli [55]. The link between the inhibition of Pavlovian responses and the balance between model-based and model-free learning may not be so surprising: Pavlovian responses are analogous to habitual ones in that both are stimulus-triggered responses, although the way they are shaped by learning differs. In particular, Pavlovian responses are innate responses to primary stimuli, such as salivation in the presence of food or freezing in the present of threat. Following learning, these responses come to be triggered anticipatorily by stimuli that are predictive of the primary stimuli, such as a bell announcing food or threat. In principle, arbitration between goal-directed and habitual responding could naturally extend to arbitration between goal-directed and Pavlovian responding [55] where the same additional costs in terms of time and executive function are required to override the default response when it may be maladaptive [78–81]. This could give a normative flavor to the hitherto more descriptive computational models of Pavlovian/instrumental competition that have been proposed [55,80,82].

Fatigue and Cognitive Control

Some theories of fatigue, defined as unwanted changes in performance owing to continued activity, propose that the sensation of fatigue indicates rising conflict between current and competing goals, and therefore tracks opportunity costs of executive resource use [56,57]. A related proposal views self-control erosion as stemming from a similar shift of motivation [83,84] that could also cleanly map to an increase in the estimated cost of the resource. However, the timescales involved are very different (hours for fatigue [85] versus minutes for self-control tasks [48]) and the motivational effects of fatigue and self-control exertion may be independent [86].

Here again, the main puzzle that needs to be addressed is why either the opportunity cost of executive resources should increase with time on task, or the benefit of resources should decrease.

Estimating the Benefit of Control: Controllability and Learned Helplessness

Early work made a distinction between 'resource-limited' and 'data-limited' response functions to resource investment, according to whether a task responds favorably to increased resource allocation [31,87]. In the data-limited regime, no resource should be applied, regardless of its current cost. How can the agent estimate the return on resource investment – that is, the value of control [3]?

As we suggested, [4,28,30] this could be identified by estimating the amount of controllability in the environment. In fact, controllability and a closely related concept, autonomy, have been shown to improve self-control performance [15] and to mitigate the worsening of self-control performance after performing a demanding task [88–90]. This effect is strikingly reminiscent of early findings that the perception of controllability alleviates the deleterious effect of stressors on performance [91,92], and these considerations are consistent with the hypothesis that the expected benefit of investing more resources into control is, at least in part, estimated by assessing the controllability of the environment [54].

Finally, controllability is most typically associated with literature on 'learned helplessness' experiments, where exposure to uncontrollable outcomes causes various degradations in subsequent decisions [58]. Learned helplessness has long been proposed as an animal model of depression; it has also recently been suggested that the symptoms of actual depression might arise, in part, from abnormal estimates of meta-decision variables such as those we consider here, which, via biasing decision control, could in turn produce characteristic disordered behavior such as anergia [93].

Learned helplessness also bears a resemblance to ego depletion in that both are experience-dependent changes in controlled behavior that generalize broadly across tasks. Indeed, many of the tasks used to probe learned helplessness and the ego depletion effect on self-control are similar – for example, measuring persistence on problem-solving or Stroop tasks [48,58,92] – potentially suggesting a common mechanism. Strengthening the contention that controllability may be a shared mechanism, correlations have been observed between emotion regulation, executive function, impulse control, and perceptions of controllability [94]; neuroimaging has shown that prefrontal activity (presumably reflecting executive function) can be modulated by controllability [95,96]; and perceptions of controllability may explain variability in the exploration–exploitation tradeoff [97].

Concluding Remarks

Many decision contexts involve a higher-order automaticity tradeoff that requires deciding what amount of resources to invest into the decision itself. We have suggested that analogous considerations arise in several different experimental circumstances broadly related to self-control, and that emerging theories and results in many of these areas reflect strikingly similar mechanisms for rational cost–benefit tradeoffs, although important outstanding questions remain (see Outstanding Questions). Similar ideas may shed light beyond the resolution of particular self-control situations, to why they arise in the first place. In particular, one can extend the rational analysis up a further level to meta-meta-decisions about the nature of controllers. These would entail decisions such as determining under what circumstances to program an automatic action (and which one), because these can also be optimized. Pavlovian and habitual responses could be the outcome of such higher-level optimizations, in one case via evolution and the other via learning over days or weeks. Exploring this even higher level of decision could yield further insights into the vast repertoire of observed decision processes and the control principles governing their interactions.

Outstanding Questions

Why should the perceived costs or benefits change over the course of self-control experiments, giving rise to fatigue, learned helplessness, or ego depletion effects?

How is the expected reward differential between competing controllers estimated?

Why is there a serial limited-capacity executive?

Acknowledgments

This work was partially supported by a Junior Fellow award from the Simons Foundation (Y.L.B.), an award in understanding human cognition from the James S. McDonnell Foundation (N.D.), and support from Google DeepMind (N.D.). We thank Peter Dayan for helpful discussions.

References

1. Kurzban, R. *et al.* (2013) An opportunity cost model of subjective effort and task performance. *Behav. Brain Sci.* 36, 661–679
2. Westbrook, A. and Braver, T.S. (2015) Cognitive effort: a neuroeconomic approach. *Cogn. Affect. Behav. Neurosci.* 15, 395–415
3. Shenhav, A. *et al.* (2013) The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron* 79, 217–240
4. Keramati, M. *et al.* (2011) Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS Comput. Biol.* 7, e1002055
5. Lieder, F. and Griffiths, T.L. (2015) When to use which heuristic: a rational solution to the strategy selection problem. In *Proceedings of the 37th Annual Conference of the Cognitive Science Society*. pp. 1362–1367, Cognitive Science Society
6. McGuire, J.T. and Botvinick, M.M. (2010) Prefrontal cortex, cognitive control, and the registration of decision costs. *Proc. Natl. Acad. Sci. U.S.A.* 107, 7922–7926
7. Kool, W. *et al.* (2010) Decision making and the avoidance of cognitive demand. *J. Exp. Psychol. Gen.* 139, 665–682
8. Dixon, M.L. and Christoff, K. (2012) The decision to engage cognitive control is driven by expected reward-value: neural and behavioral evidence. *PLoS ONE* 7, e51637
9. Kool, W. and Botvinick, M. (2014) A labor/leisure tradeoff in cognitive control. *J. Exp. Psychol. Gen.* 143, 131–141
10. Dayan, P. (2012) How to set the switches on this thing. *Curr. Opin. Neurobiol.* 22, 1068–1074
11. Kahneman, D. (2011) *Thinking, Fast and Slow,* Farrar, Straus and Giroux
12. Gershman, S. and Wilson, R. (2010) The neural costs of optimal control. In *Advances in Neural Information Processing Systems* (23) (Lafferty, J.D. *et al.*, eds), In pp. 712–720, Neural Information Processing Systems Foundation
13. Holmes, P. and Cohen, J.D. (2014) Optimality and some of its discontents: successes and shortcomings of existing models for binary decisions. *Top. Cogn. Sci.* 6, 258–278
14. Botvinick, M.M. and Cohen, J.D. (2014) The Computational and Neural Basis of Cognitive Control: Charted Territory and New Frontiers. *Cogn. Sci.* 38, 1249–1285
15. Kanfer, F.H. and Seidner, M.L. (1973) Self-control: factors enhancing tolerance of noxious stimulation. *J. Pers. Soc. Psychol.* 25, 381–389
16. Payne, J.W. *et al.* (1988) Adaptive strategy selection in decision making. *J. Exp. Psychol. Learn. Mem. Cogn.* 14, 534–552
17. Gluth, S. *et al.* (2012) Deciding when to decide: time-variant sequential sampling models explain the emergence of value-based decisions in the human brain. *J. Neurosci.* 32, 10686–10698
18. Hofmann, W. *et al.* (2009) Three ways to resist temptation: the independent contributions of executive attention, inhibitory control, and affect regulation to the impulse control of eating behavior. *J. Exp. Soc. Psychol.* 45, 431–435
19. Etzel, J.A. *et al.* (2015) Reward motivation enhances task coding in frontoparietal cortex. *Cereb. Cortex* Published online January 19, 2015. http://dx.doi.org/10.1093/cercor/bhu327
20. Job, V. *et al.* (2010) Ego depletion – is it all in your head? Implicit theories about willpower affect self-regulation. *Psychol. Sci.* 21, 1686–1693
21. Gross, J.J. and Thompson, R.A. (2007) Emotion regulation: conceptual foundations. In *Handbook of Emotion Regulation* (Gross, J.J., ed.), pp. 3–24, Guilford Press
22. Daw, N.D. *et al.* (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8, 1704–1711
23. Pezzulo, G. *et al.* (2013) The mixed instrumental controller: using value of information to combine habitual choice and mental simulation. *Front. Psychol.* 4, 92
24. Bogacz, R. *et al.* (2006) The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol. Rev.* 113, 700–765
25. Dehaene, S. *et al.* (1998) A neuronal model of a global workspace in effortful cognitive tasks. *Proc. Natl. Acad. Sci. U.S.A.* 95, 14529–14534
26. Baum, E.B. and Smith, W.D. (1997) A Bayesian approach to relevance in game playing. *Artif. Intell.* 97, 195–242
27. Koechlin, E. and Summerfield, C. (2007) An information theoretical approach to prefrontal executive function. *Trends Cogn. Sci.* 11, 229–235
28. Gittins, J. (1989) *Multi-Armed Bandit Allocation Indices,* John Wiley & Sons
29. Dayan, P. (2013) Exploration from generalization mediated by multiple controllers. In *Intrinsically Motivated Learning in Natural and Artificial Systems* (Baldassarre, G. and Mirolli, M., eds), pp. 73–91, Springer
30. Wilson, R.C. *et al.* (2014) Humans use directed and random exploration to solve the explore–exploit dilemma. *J. Exp. Psychol. Gen.* 143, 2074–2081
31. Norman, D.A. and Bobrow, D.G. (1976) On the analysis of performance operating characteristics. *Psychol. Rev.* 83, 508–510
32. Ryan, R.M. and Deci, E.L. (2006) Self-regulation and the problem of human autonomy: does psychology need choice, self-determination, and will? *J. Pers.* 74, 1557–1585
33. Tversky, A. and Simonson, I. (1993) Context-dependent preferences. *Manag. Sci.* 39, 1179–1189
34. Morgan, K.V. *et al.* (2012) Context-dependent decisions among options varying in a single dimension. *Behav. Processes* 89, 115–120
35. Vasconcelos, M. *et al.* (2013) Context-dependent preferences in starlings: linking ecology, foraging and choice. *PLoS ONE* 8, e64934
36. Huang, H.J. and Ahmed, A.A. (2014) Older adults learn less, but still reduce metabolic cost, during motor adaptation. *J. Neurophysiol.* 111, 135–144
37. Dayan, P. (2014) Rationalizable Irrationalities of Choice. *Top. Cogn. Sci.* 6, 204–228
38. Stephens, D.W. and Krebs, J.R. (1986) *Foraging Theory,* Princeton University Press
39. Niv, Y. *et al.* (2007) Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology* 191, 507–520
40. Kacelnik, A. (1997) Normative and descriptive models of decision making: time discounting and risk sensitivity. *Ciba Found. Symp.* 208, 51–67
41. Cools, R. *et al.* (2011) Serotonin and dopamine: unifying affective, activational, and decision functions. *Neuropsychopharmacology* 36, 98–113
42. Dezfouli, A. and Balleine, B.W. (2012) Habits, action sequences and reinforcement learning. *Eur. J. Neurosci.* 35, 1036–1051
43. Mani, A. *et al.* (2013) Poverty impedes cognitive function. *Science* 341, 976–980
44. Baddeley, A. (2012) Working memory: theories, models, and controversies. *Annu. Rev. Psychol.* 63, 1–29
45. Hofmann, W. *et al.* (2012) Executive functions and self-regulation. *Trends Cogn. Sci.* 16, 174–180
46. Raio, C.M. *et al.* (2013) Cognitive emotion regulation fails the stress test. *Proc. Natl. Acad. Sci. U.S.A.* 110, 15139–15144
47. Otto, A.R. *et al.* (2013) Working-memory capacity protects model-based learning from stress. *Proc. Natl. Acad. Sci. U.S.A.* 110, 20941–20946
48. Hagger, M.S. *et al.* (2010) Ego depletion and the strength model of self-control: a meta-analysis. *Psychol. Bull.* 136, 495–525
49. Schmeichel, B.J. (2007) Attention control, memory updating, and emotion regulation temporarily reduce the capacity for executive control. *J. Exp. Psychol. Gen.* 136, 241–255
50. Otto, A.R. *et al.* (2015) Cognitive control predicts use of model-based reinforcement learning. *J. Cogn. Neurosci.* 27, 319–333
51. Schmeichel, B.J. *et al.* (2008) Working memory capacity and the self-regulation of emotional expression and experience. *J. Pers. Soc. Psychol.* 95, 1526–1540

52. Krebs, J.R. *et al.* (1977) Optimal prey selection in the great tit (*Parus major*). *Anim. Behav.* 25, 30–38
53. Krebs, J.R. *et al.* (1974) Hunting by expectation or optimal foraging? A study of patch use by chickadees. *Anim. Behav.* 22, 953–964
54. Huys, Q.J.M. and Dayan, P. (2009) A Bayesian formulation of behavioral control. *Cognition* 113, 314–328
55. Dayan, P. *et al.* (2006) The misbehavior of value and the discipline of the will. *Neural Netw.* 19, 1153–1160
56. Hockey, R. (2013) *The Psychology of Fatigue: Work, Effort and Control,* Cambridge University Press
57. Hockey, G.R.J. (2011) A motivational control theory of cognitive fatigue. In *Cognitive Fatigue: Multidisciplinary Perspectives on Current Research and Future Applications* (Ackerman, P.L., ed.), pp. 167–187, American Psychological Association
58. Miller, W.R. and Seligman, M.E. (1975) Depression and learned helplessness in man. *J. Abnorm. Psychol.* 84, 228–238
59. Dolan, R.J. and Dayan, P. (2013) Goals and habits in the Brain. *Neuron* 80, 312–325
60. Gillan, C.M. *et al.* (2015) Model-based learning protects against forming habits. *Cogn. Affect. Behav. Neurosci.* 15, 523–536
61. Friedel, E. *et al.* (2014) Devaluation and sequential decisions: linking goal-directed and model-based behavior. *Front. Hum. Neurosci.* 8, 587
62. Daw, N.D. and Dayan, P. (2014) The algorithmic anatomy of model-based evaluation. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 369, 20130478
63. Armel, K.C. and Rangel, A. (2008) *The Impact of Computation Time and Experience on Decision Values,* American Economic Association
64. Armel, K.C. *et al.* (2008) *Biasing Simple Choices by Manipulating Relative Visual Attention,* Society for Judgment and Decision Making
65. Sokol-Hessner, P. *et al.* (2012) Decision value computation in DLPFC and VMPFC adjusts to the available decision time. *Eur. J. Neurosci.* 35, 1065–1074
66. Liljeholm, M. *et al.* (2013) Neural correlates of the divergence of instrumental probability distributions. *J. Neurosci.* 33, 12519–12527
67. Robinson, M.D. *et al.* (2010) A cognitive control perspective of self-control strength and its depletion. *Soc. Personal. Psychol. Compass* 4, 189–200
68. Baumeister, R.F. *et al.* (2007) The strength model of self-control. *Curr. Dir. Psychol. Sci.* 16, 351–355
69. Vohs, K.D. *et al.* (2012) Motivation, personal beliefs, and limited resources all contribute to self-control. *J. Exp. Soc. Psychol.* 48, 943–947
70. Eisenberger, R. (1992) Learned industriousness. *Psychol. Rev.* 99, 248–267
71. Bandura, A. (1997) *Self-Efficacy: The Exercise of Control,* W.H. Freeman/Times Books/Henry Holt & Co
72. Gist, M.E. and Mitchell, T.R. (1992) Self-efficacy: a theoretical analysis of its determinants and malleability. *Acad. Manage. Rev.* 17, 183–211
73. Eisenberg, N. *et al.* (2007) Effortful control and its socioemotional consequences. In *Handbook of Emotion Regulation* (Gross, J.J., ed.), pp. 287–306, Guilford Press
74. Hartley, C.A. and Phelps, E.A. (2009) Changing fear: the neurocircuitry of emotion regulation. *Neuropsychopharmacology* 35, 136–146
75. Ochsner, K.N. *et al.* (2012) Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion. *Ann. N. Y. Acad. Sci.* 1251, E1–E24
76. Goldin, P.R. *et al.* (2008) The neural bases of emotion regulation: reappraisal and suppression of negative emotion. *Biol. Psychiatry* 63, 577–586
77. Bush, G. *et al.* (2000) Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn. Sci.* 4, 215–222
78. Rigoli, F. *et al.* (2012) Aversive Pavlovian responses affect human instrumental motor performance. *Front. Neurosci.* 6, 134
79. Hershberger, W.A. (1986) An approach through the looking-glass. *Anim. Learn. Behav.* 14, 443–451
80. Lesaint, F. *et al.* (2014) Accounting for negative automaintenance in pigeons: a dual learning systems approach and factored representations. *PLoS ONE* 9, e111050
81. Cavanagh, J.F. *et al.* (2013) Frontal theta overrides Pavlovian learning biases. *J. Neurosci.* 33, 8541–8548
82. Lesaint, F. *et al.* (2014) Modelling individual differences in the form of Pavlovian conditioned approach responses: a dual learning systems approach with factored representations. *PLoS Comput. Biol.* 10, e1003466
83. Inzlicht, M. *et al.* (2014) Why self-control seems (but may not be) limited. *Trends Cogn. Sci.* 18, 127–133
84. Huizenga, H.M. *et al.* (2012) Muscle or motivation? A stop-signal study on the effects of sequential cognitive control. *Front. Psychol.* 3, 126
85. Mackworth, N.H. (1948) The breakdown of vigilance during prolonged visual search. *Q. J. Exp. Psychol.* 1, 6–21
86. Vohs, K.D. *et al.* (2011) Ego depletion is not just fatigue: evidence from a total sleep deprivation experiment. *Soc. Psychol. Personal. Sci.* 2, 166–173
87. Norman, D.A. and Bobrow, D.G. (1975) On data-limited and resource-limited processes. *Cognit. Psychol.* 7, 44–64
88. Muraven, M. *et al.* (2008) Helpful self-control: autonomy support, vitality, and depletion. *J. Exp. Soc. Psychol.* 44, 573–585
89. Moller, A.C. *et al.* (2006) Choice and ego-depletion: the moderating role of autonomy. *Pers. Soc. Psychol. Bull.* 32, 1024–1036
90. Muraven, M. *et al.* (2007) Lack of autonomy and self-control: performance contingent rewards lead to greater depletion. *Motiv. Emot.* 31, 322–330
91. Glass, D.C. *et al.* (1969) Psychic cost of adaptation to an environmental stressor. *J. Pers. Soc. Psychol.* 12, 200–210
92. Glass, D.C. *et al.* (1973) Perceived control of aversive stimulation and the reduction of stress responses. *J. Pers.* 41, 577–595
93. Huys, Q.J.M. *et al.* (2015) Depression: a decision-theoretic analysis. *Annu. Rev. Neurosci.* 38, 1–23
94. Declerck, C.H. *et al.* (2006) On feeling in control: a biological theory for individual differences in control perception. *Brain Cogn.* 62, 143–176
95. Xue, G. *et al.* (2013) Agency modulates the lateral and medial prefrontal cortex responses in belief-based decision making. *PLoS ONE* 8, e65274
96. Walton, M.E. *et al.* (2004) Interactions between decision making and performance monitoring within prefrontal cortex. *Nat. Neurosci.* 7, 1259–1265
97. Kayser, A.S. *et al.* (2015) Dopamine, locus of control, and the exploration–exploitation tradeoff. *Neuropsychopharmacology* 40, 454–462
98. Norman, D.A. and Shallice, T. (1986) Attention to action. In *Consciousness and Self-Regulation* (Davidson, R.J. *et al.*, eds), pp. 1–18, Springer
99. Granon, S. and Changeux, J-P. (2012) Deciding between conflicting motivations: what mice make of their prefrontal cortex. *Behav. Brain Res.* 229, 419–426
100. Koechlin, E. (2014) An evolutionary computational theory of prefrontal executive function in decision-making. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 369, 20130474