Trends in
Cognitive Sciences

CellPress

Review

Why is cognitive effort experienced as costly?

A. Ross Otto [1,]*, Andrew Westbrook[2], and Jean Daunizeau[3]

A widespread observation is that people avoid mentally effortful courses of action, and much recent work examining cognitive effort has explained subjective effort evaluation – and, consequently, preferences – in economic terms, which assumes that the expenditure of cognitive effort is experienced as costly. However, this economic perspective is largely tacit about the source of these costs. Here, we review recent theoretical treatments of effort costs, which take vastly different perspectives (information-theoretic, psychological, and biological) to explain how the subjective experience of cognitive effort arises from controlled information processing, exploring their predictions concerning the simple observation that people experience tasks with high (versus low) working memory demands as costly. Finally, we identify open questions that might help bridge across these accounts.

Highlights

A body of research suggests that people often shy away from mentally demanding tasks, with many studies explaining this behavior through an economic lens that views cognitive effort as inherently unpleasant.

While influential economic accounts assume that effort carries a subjective cost, it remains unclear how this cost might arise from first principles.

This review critically examines diverse theoretical perspectives – spanning information theory, psychology, and network control theory – which attempt to explain how controlled information processing generates the subjective experience of mental effort.

The review ultimately seeks to identify key unresolved questions that could help synthesize these varied theoretical approaches, providing a more comprehensive understanding of how and why people experience certain forms of information processing as effortful.

The subjective experience of cognitive effort

Why does the exertion of certain forms of cognitive effort feel costly? A near-universal assumption underlying the study of human cognition is that a variety of mental activities are constrained by limitations in human information processing capacity [1–3]. While earlier endeavors assumed different structural constraints on information processing [4], they agree that many volitional mental activities (e.g., performing mental arithmetic) are experienced as effortful, but that these constraints do not seem to apply to forms of information processing understood to occur automatically, rapidly, and unconsciously (e.g., early processing of visual input or control of balance and posture) [5,6]. Here, we review theories postulating why we often experience controlled – versus automatic [7] – information processing as effortful and unpleasant.

A common observation in cognitive effort research is that individuals avoid activities that are experienced as effortful [8,9], a conjecture first made over a century ago, referred to as the ‘law of least effort’ [10,11]. This preference against mentally effortful activities has been interpreted, following economic orthodoxy [12,13], as an indication that the experience of cognitive effort carries negative utility [8,14] and can evoke negative affect [15,16]. Based on these two observations, it stands to reason that people typically evaluate the exertion of cognitive effort as ‘costly’ (but see [17–20]).

Here, we focus on an important, but less well understood question: why is cognitive effort exertion costly? We center on understanding principles giving rise to the costly (subjective) experience of effort rather than on understanding the factors governing effort exertion, typically assessed via task performance or physiological measures, such as heart rate and pupillometry [21,22]. While it appears obvious that the brain must somehow compute effort costs to produce the subjective experience of mental effort exertion, cognitive neuroscience lacks a principled, consensus understanding of the source of cognitive effort costs. Yet, several intuitively appealing theories have been proposed that attempt to explain the source of effort costs at different levels of analysis: neurobiological, computational, and psychological.

[1]Department of Psychology, McGill University, Montreal, Quebec, Canada
[2]Department of Psychiatry, Rutgers University, Piscataway, NJ, USA
[3]Paris Brain Institute, Sorbonne Université/INSERM UMR S1127, Paris, France

*Correspondence: ross.otto@mcgill.ca (A.R. Otto).

Quantifying subjective effort costs

It is important to distinguish between the perception versus the exertion of effort [23,24], which are, to some extent, dissociable [21,25]. With respect to the latter, previous theoretical frameworks have endeavored to explain how the costs and benefits of effort exertion guide cognitive investment (Box 1). Critically, rather than posit a source of effort costs, these accounts presuppose that cognitive effort is costly and, thus, its exertion should be justified by the benefits it confers. Empirically, subjective effort costs manifest, most notably, in self-reported aversiveness ratings, and in preferences for less demanding courses of action [8,14]. Effort costs and experienced task difficulty are also closely related but not identical constructs: a task may be difficult but not necessarily effortful – for example, perception of a visually degraded stimulus [9].

Here, we explore three broad frameworks, namely, processing bottlenecks, information theory, and brain network control, motivating our review with findings from a classic working memory paradigm. In the *N*-back task (Figure 1A), participants are presented with a stream of stimuli and indicate whether each stimulus matches the stimulus *N* items before ('target') or not ('nontarget') [26]. Load increases with larger *N* values, possibly reflecting increased demands on monitoring, maintenance, inhibition, updating, and interference resolution processes [27–29]. In the context of effort research, the *N*-back task has yielded several convergent findings, suggesting that increases in (objective) task demand level produce corresponding increases in subjective effort costs: first, self-reported effort increases with *N*-back level [30,31]. Second, given the choice between two levels (e.g., a 1- versus a 3-back), individuals typically indicate a preference for the less demanding option [30]. Third, participants forego money and will even elect to receive physical pain to avoid performing the most difficult *N*-back levels [31,32].

With this empirical characterization of subjective effort in the *N*-back task, we next explore how these recent theoretical accounts – each making a specific commitment to a source of effort costs – can explain, from first principles, how parametric increases in *N* level might give rise to observed preferences and self-reported effort (Figure 1B,C). We also note that a diversity of paradigms can measure subjective effort costs. For example, patch-foraging tasks can be used to evaluate effort costs in terms of the degree to which participants' decide to 'travel'

Box 1. Accounts of cognitive effort exertion

According to one influential perspective rooted in economics, the intensity with which an individual exerts cognitive effort – or, alternatively, their effort investment level – is governed by a simple cost–benefit trade-off. Simply put, effortful mental processing should be engaged to the extent that the benefits of effortful mental processing activity outweigh the costs they incur [14,76,77]. In support of this cost–benefit view, a body of empirical work supports the idea that reward incentives in effect 'offset' the costs of cognitive effort, finding that increased reward incentives promote intensification of effortful cognitive control as measured by task performance relative to circumstances when reward incentives are smaller or absent [8,78–80]. Likewise, faced with explicit choices about effort expenditure, individuals are more willing to engage in effortful mental activity when larger reward incentives are on offer, but are less willing to engage in tasks that impose greater task demands (and involve greater subjective effort) [81–84].

A related and influential framework – Motivational Intensity Theory [85,86] – posits that individuals invest effort in a given task jointly in accordance with perceived task demand (i.e., difficulty) and the importance of task success. On this view, individuals seek to avoid 'wasted' effort: task success must be important enough to justify expending the effort required for a task of a given difficulty level. Supporting this view, a considerable body of empirical work suggests that effort exertion, indexed by cardiovascular reactivity, is maximal when effort is justified and task difficulty is high (but not impossible) [19,85,87,88]. Similarly, studies examining task performance observe that when success importance is low (e.g., when increased effort exertion is unlikely to confer additional rewards), individuals are less willing to increase effort exertion in accordance with reward incentives [78,89,90].

However, rather than postulating a source of subjective effort costs or making a commitment to a mechanism by which these effort costs arise, these frameworks are chiefly concerned with an individual's level of effort exertion, taking subjective (or perceived) effort as an input and a cost to be justified.

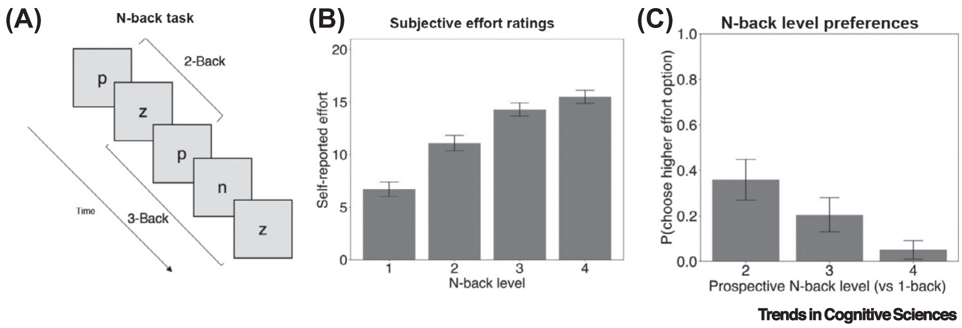

Figure 1. The *N*-back task. (A) In this paradigm, participants view a sequence of letter stimuli presented one at a time and must decide whether each stimulus matches the one presented *N* steps back in the sequence. (B) Subjective effort ratings as a function of working demand (*N* level) in the *N*-back task. (C) Preferences for different working memory demand (*N*) levels when paired with the option of performing a 1-back task. Data replotted from [31] (B) and [30] (C).

(which incurs cognitive demand) versus 'stay' [33,34]. In other paradigms, effort costs have been quantified in terms of implicit effort discounting, whereby participants 'discount' rewards earned while performing demanding tasks [35–37]. However, one limitation inherent in many of these empirical approaches is that our subjective evaluation of effort is not, in many cases, a direct reflection of 'objective' task demand. For example, both self-report ratings and choices concerning task demand levels appear systematically biased by the context in which effortful activity is undertaken [38] and options are evaluated [39–41], which suggests that subjective effort signals are not a simple readout of objective demand, but incorporate other experiential factors.

Architectural constraints and processing bottlenecks

By one perspective, cognitive control (i.e., the capacity to override habitual and automatic responses in favor of arbitrary, goal-congruent thought and action [42]) is costly because of bottlenecks in control architecture. Given that cognitive control relies upon shared and limited processing resources, allocating these resources to one task necessarily implies missing out on other opportunities. In other words, prioritizing resources for a given goal entails an opportunity cost (Box 2). Importantly, accounting for this cost is adaptive, because it helps ensure that

Box 2. Opportunity cost accounts of cognitive effort

A related economics-inspired theory of effort costs proposes that subjective effort costs arise from the opportunity costs of occupying (limited-capacity) cognitive processing in the service of one particular task; that is, the possible benefits that one foregoes by engaging these same resources for some other, alternative task [91]. On this view, the cost of engaging a processing resource in one task is computed as the expected utility conferred by the next-best possible alternative task (to which these same processing resources could be allocated) [92]. Indeed, this notion of a cost follows naturally from serial processing constraints imposed by the structural bottleneck view of information processing [47], and may enable individuals to efficiently prioritize processing demands among competing goals/tasks. In support of this idea, recent studies observed that fluctuations in effortful cognitive processing [93,94] as well as self-reported levels of fatigue and boredom [95,96], can be explained, in part, by opportunity costs, operationalized in diverse ways.

How might an opportunity costs account explain the increase in subjective effort accompanying increasing working memory demands in the *N*-back task (Figure 1 in the main text)? Under the assumption that, for an individual performing the *N*-back task, the expected utility of the next-best alternative task (e.g., mind-wandering) in which that individual could plausibly engage does not vary as a function of *N* level, it is unclear why subjective effort would increase between a 2- and 3-back task. Rather, since the value of the next-best use of computational processes is invariant across *N* levels, opportunity costs – and thus subjective effort – should remain constant across *N* levels. By contrast, if one assumes that some processing resources remain available for mind-wandering during the *N*-back task, but that their availability decreases as *N* level increases, then the additional opportunity cost incurred in the 3-back (relative to the 2-back) could plausibly generate subjective effort costs that scale with *N*. In short, the ability of this account to explain subjective effort patterns hinges critically on a specific definition of opportunity costs, which we contend may be under-constrained.

resources will be reserved for tasks with the highest expected value [14]. Under this view, cognitive effort costs are effectively opportunity costs that arise from 'bottlenecked' resources. But why are there processing bottlenecks in the first place? If control is indispensable, why would the brain not be equipped with multiple, independent control modules to augment its information-processing capacity? Complete independence among control systems could result in disorganized behaviors, competing for time and energy, thus motivating the notion of a central executive [43] capable of prioritizing across tasks. Moreover, just as parallel motor control signals inevitably interfere at the level of motor effectors, independent cognitive control signals would interfere at the level of bottom-up cognition (an analogy motivated, in part, by the common algorithms and operational principles of cognitive and motor control [44]). While complete independence among control systems is impractical, in principle, one can imagine a hierarchy involving a unitary prioritization mechanism and multiple, parallel control modules devoted to separate tasks. Yet, in reality, multitasking (i.e., executing multiple tasks and operations in parallel) is only possible with extreme amounts of practice and, even then, is limited to well-practiced tasks with mostly orthogonal stimulus and response modalities [45]. Thus, these limitations in multi-tasking suggest shared and limited processing resources [46].

Although bottlenecked resources may not be optimal for multitasking, they do offer computational advantages. Namely, through multiplexing – whereby task representations are shared across multiple, distinct tasks – we are better at learning and generalization [47]. A key parameter that may determine the capacity for learning – and thus the relative advantage of bottlenecked resources – is the dimensionality of task representations, where low- versus high-dimensional task representations are better for generalization. Importantly, this parameter may have implications for effort costs as well, which we explore in Box 3.

Box 3. Dimensionality of control representations and effort costs

One architectural factor that may bear on the subjective cost of cognitive effort is the dimensionality of control representations. Here, 'control representations' refer to neural ensembles that convey control signals to bias goal-directed perception and action selection, and 'dimensionality' is defined as the number of orthogonal axes needed to describe the variance in ensemble activity across contexts. Dimensionality may bear on the costliness of control, given that it mediates a trade-off between separability and generalizability [97].

The separability–generalizability trade-off refers to a continuum between high precision mappings of inputs to outputs (separability) and broad generalization across variable and even noisy inputs (generalizability; see [97] for a comprehensive review). High-dimensional representations support separability by making fine distinctions among inputs, parsing even subtly different states into precise responses. This capacity for fine distinctions is essential for selecting and implementing action policies, the complexity of which appears to trade off against reward [98]. High-dimensional task representations are also essential for performing complex tasks well [99,100]. By contrast, the downside of highly resolved input–output mappings is that they are also brittle: even a small degree of variability or noise in inputs can result in dramatically different outputs. Conversely, low-dimensional representations are more likely to generate consistent outputs in the face of variable and noisy inputs. By the same token, low-dimensional representations are more capable of leveraging similarities across tasks, thus facilitating generalization and learning. Therefore, low-dimensional control representations may be more useful early, when new tasks need to be learned.

These distinctions may have implications for subjective effort – indeed, subjective effort costs have been proposed as a solution to the problem of high dimensionality in the solution space [101]. For example, high-dimensional representations may meliorate opportunity costs by facilitating flexible task-switching in multitask contexts. This capacity to facilitate task switching derives from the fact that small differences in high-dimensional control representations can result in substantially distinct outputs [97]. Thus, with well-practiced tasks, switching requires minimal reconfiguration when relying on high-dimensional control representations. This might be especially useful, for example, in stemming the inflation of subjective effort with higher *N*-back load levels involving a greater number of subtasks in each trial. Conversely, there is evidence that high-dimensional representations, and their capacity for separability, require a higher coding level, referring to the average number of neurons engaged by task representations [102]. Thus, high-dimensional representations may impose greater metabolic costs with potential implications for subjective effort (see Box 4 in the main text).

A key downside of multiplexed resources is that multiple, concurrent demands can cause crosstalk interference if tasks that share representations are executed simultaneously. Given that even small degrees of overlap in shared representations can result in catastrophic interference [48], it is essential to restrict ourselves to one task at a time. Thus, bottlenecking may reflect an adaptive strategy for systems working with shared representations, which are otherwise beneficial for learning and generalization, rather than an inherent limitation on control system bandwidth. In this context, a central purpose of cognitive control may be to prioritize competing solicitations of shared resources to avoid debilitating crosstalk [47].

Why then, according to the bottlenecking hypothesis, do higher *N*-back levels incur larger effort costs? Prioritization of selected tasks deprioritizes others, thus giving rise to opportunity costs (Box 2). We speculate that effort costs of the 2-back are higher than the 1-back because the degree (or density) of bottlenecking per unit time (or trial) is greater, thus increasing opportunity costs. Note that the 2-back involves a larger number of working memory subtasks (encoding, maintenance, comparison, positional updates, inhibition, etc.) in each trial relative to the 1-back (encoding, maintenance, and comparison). Thus, working memory faculties that might otherwise be allocated to mind-wandering are even more sharply limited in the case of the 2-back task, where bottlenecking density is higher, increasing opportunity costs.

The bottlenecking account makes at least two predictions for which there is already empirical support. First, larger opportunity costs accumulate the longer that bottlenecks are imposed. Thus, effort costs should bias against the persistence of task representations. Indeed, given the choice between working memory tasks involving more frequent updating or more persistent maintenance of working memory contents, individuals prefer more frequent updating [49], suggesting that persistent representations are experienced as more costly. Second, tasks with more overlapping processes will be costlier because more intensive bottlenecks are required to limit processing overlap. This might arise, for example, in task-switching paradigms where stimuli and responses are identical, but their mapping depends on rule switches. Indeed, a body of work finds that individuals prefer to avoid task switching [8,50]. Similarly, the potential for distractor interference with similar working memory contents in the *N*-back (e.g., so-called 'lures', which are close to, but not in the exact sequential position to be, a target [51]) may explain why higher *N* levels are costlier. Indeed, a recent experiment quantifying the relative cost of subprocesses – by examining wages demanded by participants to perform subtly distinct *N*-back task designs – observed that distractor resistance is one of the more effort-costly subprocesses in the *N*-back [29]. This perspective is consistent with the hypothesis that working memory capacity is limited, primarily, by interference from distractors [52]. Thus, bottlenecking intensity and effort costs should increase as a function of the potential for interference.

Information-theoretic view of effort costs

Under a recent information-theoretic account [53], effort costs arise during inference and/or action selection as a function of the amount of information required for an individual to update their prior beliefs (e.g., probabilities over possible responses) to posterior beliefs (e.g., response probabilities conditioned on perceptual input). In information-theoretic terms, the computational cost of an inference (or action) is proportional to the reduction in relative entropy accompanying the update of a prior belief distribution to a posterior distribution, which can be quantified in terms of the Kullback–Leibler (KL) divergence between these two distributions. Put simply, inferences and actions engendering larger informational costs incur larger effort costs. The idea that unlikely actions deviating from one's 'default' policy, are treated as costly because they require costly planning (i.e., inference) was put forward in an influential proposal [54]. Relatedly, recent empirical work suggests that mutual information between an individual's actions and outcomes heightens the subjective experience of 'flow' or 'immersion' in a task [55], possibly owing to feelings of control [56].

A key prediction of this theory is that tasks with many response options characteristically entail a widely spread prior belief distribution over possible responses, leading to a large difference between the prior and posterior belief distributions. In turn, the increased informational cost of selecting among many possible responses engenders additional cognitive costs (evinced by slower response thresholds) of considering additional (*a priori* equally likely) response options. Considering the patterns of subjective effort observed in the *N*-back (Figure 1B), increased informational costs might arise as a function of working memory demand *N* stemming from the divergence between the prior distribution (governed by the number of possible responses, which is constant) and the size of the posterior distribution (defined by the sequence of *N* recent encoded items, which scales with *N*). Owing to the reduction in uncertainty (i.e., KL divergence between prior and posterior) being larger in the case of the 2-back versus the 1-back, this might lead to larger effort costs as *N* increases. From this perspective, the increasing demands that *N* levels impose upon specific working memory processes (e.g., storing, binding, and updating with respect to sequence position [27]) are captured by these information costs.

A brain network perspective on the cost of cognitive control

While much of the energetic budget of the brain is fixed, a significant portion is spent on operating costs that adapt to cognitive demand (Box 4). Fluctuating demands may especially overwhelm the capacity of the automatic processing systems of the brain to adapt, hence calling on cognitive control to flexibly adjust the intrinsic dynamics of the brain. In particular, flexibility would be required for overriding automatic (i.e., reflexive or habitual) responses in the context of complex cognitive tasks (e.g., reasoning, planning, etc.).

One possibility is that the cost of controlled processes stems from engaging specific brain systems that support flexible, adaptive cognitive control – most notably, the prefrontal cortex

Box 4. Energy constraints: how does the brain regulate the allocation of energetic/metabolic resources?

The brain is metabolically expensive. In humans, it consumes ~20% of our metabolic energy, despite comprising only 2% of our body mass [103–105]. About 80% of the energy in the brain is used for maintaining its baseline state. This part of the energetic budget is mostly spent on either maintenance of ion gradients across neuronal membranes, synaptic maintenance, or cellular housekeeping [106–108]. The remaining 20% of the energetic budget is spent on local operational costs that fluctuate depending on cognitive demand, which essentially results from task-related electrical signaling [109]. In turn, although transient and regionally specific [110,111], the operational costs of the brain are a non-negligible part of the total energy budget of the body.

This suggests that animals have evolved adaptive strategies to maintain their expansive brains [112]. First, according to the 'Expensive Brain Hypothesis', animals balance the energy needs of the brain by reallocating energy that would otherwise be spent on other bodily functions [113]. For example, long-term memory formation in flies can impair survival during food/water shortages or immune challenges [114]. Second, the brain has evolved energy-efficient coding strategies to operate at reduced cost. For example, according to the 'efficient coding' hypothesis, the brain optimizes the information transmitted per ATP, eventually favoring neuronal firing rates that are lower than those that maximize information transfer [115,116]. Finally, the brain adapts its energetic budget allocation when environmental resources are scarce; this requires both a decrease in energy expenditure of nonessential neural functions and selective enhancement of useful neural functions. For example, although fruit flies specifically disable the costly formation of long-term memory upon starvation [117], the precision of neural representations for food cues is selectively enhanced [118,119].

But is the brain saving its energetic resources at the expense of its own information processing capacity, even when environmental resources are abundant? Recent work shows that artificially unlocking mitochondrial metabolism – the main energy supply for neurons - in central memory circuits of flies and mice eventually boosts long-term memory storage [120]. This is important because it suggests that an evolutionarily conserved mechanism sets the energy–performance balance of the brain by actively limiting mitochondrial metabolism in active neurons.

(PFC). Interestingly, although the PFC mostly shares its cellular composition and layered organization with the rest of the neocortex, its brain-wide connectivity pattern is unique [57–60]. While this may have computational consequences, it does not necessarily mean that prefrontal computations are costlier. What is lacking is quantitative insight into how brain architecture (as defined, e.g., by connectomics) determines the operational costs and hence, the subjective effort costs, associated with engagement of controlled processes.

Network control theory is a simple and powerful tool for predicting the energy required to guide a system toward a target state, given system architecture [61]. Here, the state of the system is defined as the pattern of neuronal activity across brain regions at a single time point. The target state is chosen for its functional utility (e.g., because of its information content), but may not be reached easily. In this context, the controllability of a network refers to the possibility of manipulating network nodes (using local inputs) to reach target states. Critically, the ability of a particular node (i.e., brain region) to control the state of the network depends upon the nature of the target state and the location of the node within the network.

In control-theoretic terms, automatic and controlled processes correspond to target states that are easy and difficult to reach, respectively [62]. Formally, a target state is difficult to reach because it is unstable or transient, for example, because it relies heavily on sparsely connected nodes, thus requiring specific energy allocation patterns [63]. In turn, hard-to-reach states are rare and typically require substantial input energy. This perspective is interesting because it formally relates controlled processes to effort costs without speculating about the complexity of the underlying cognitive computations. Rather, large effort costs would ensue from structural constraints imposing higher energy inputs for controlled processes.

It is worth drawing a distinction between two forms of controllability relevant for this hypothesis [62]. 'Modal controllability' measures how efficiently a given node can steer the system into hard-to-reach states [64]. Brain regions that exhibit high modal controllability are nodes that are ideally positioned in the network to reach or sustain hard-to-reach states. By contrast, 'average controllability' measures how many target states can be easily reached from a given node. Brain regions that exhibit high average controllability are those network nodes that facilitate transitions toward easy-to-reach states. Both measures can be derived from brain-wide structural connectivity using diffusion-weighted MRI [65]. Interestingly, modal controllability is greatest in the frontoparietal and cingulo-opercular regions, whereas average controllability is greatest within the default mode network [62]. This suggests that brain subsystems thought to underpin cognitive control turn out to exhibit connectivity profiles that position them best to move the brain into hard-to-reach states – albeit at a high energetic cost.

In summary, this control-theoretic perspective suggests that: (i) the structural connectivity of the brain determines whether a given target state will be easily reachable (e.g., automatic process) or not (controlled process); and (ii) controlled/effortful processes underpinned by the PFC correspond to hard-to-reach states which, by definition, require more energy to attain (Box 4). Implicit in this theory is the notion that the brain can self-monitor with perfect accuracy. Recent developments of this theory essentially aim at relaxing this assumption (Box 5).

This perspective provides an elegant explanation for why the subjective effort increases with cognitive demand in the *N*-back task (Figure 1B): more difficult tasks require accessing or sustaining target states that are relatively unstable [66], making them harder to reach and/or maintain. This is particularly relevant for working memory representations, which tend to lose stability over time [67–69]. Thus, on average, successfully completing a 2-back task likely requires sustaining

Box 5. The controllosphere

Network control theory can be used to explain how cognitive effort costs arise from simple assumptions about the impact of structural constraints on the energetic efficiency of the information-processing routes in the brain [62,64], but this perspective may ignore some perceptual/experiential aspects of cognitive effort. In particular, implicit in this reasoning is the notion that subjective effort costs arise from the ability of the brain to monitor its own operational costs, including within nodes sparsely connected to the rest of the network. This is especially important in tasks requiring precise cognitive computations (e.g., planning), because self-monitoring by the brain enables it to proactively detect and respond to internal errors or inefficiencies before they cause overt behavioral changes. However, some controlled processes appear to induce little, if any, subjective mental effort. For example, sustained attentional states associated with 'flow' – an experiential state occurring when an individual's degree of focus matches the challenge of the task at hand [121] – are not typically perceived as effortful, even in task settings that arguably rely on controlled processes. One possibility is that experienced effort may not relate to the amount of energy expenses per se, but rather to the perceived inefficiency of this outlay for achieving cognitive goals [122], referring to the monitored confidence gain per unit of invested resources [123]. In this context, cognitive effort may simply be the difference between energy expenditure and its perceived consequence – in terms of internal estimates of goal achievement – under the constraint that the self-monitoring system of the brain is partially ignorant of hard-to-reach states. Recall that any mental work requires movement between brain states. By definition, the 'controllosphere' is the region of the state space populated by hard-to-reach states [124]. In this view, a flow state is experienced as effortless because part of the underlying energy expenditure induces motion of detectable states outside the controllosphere (i.e., through easy-to-reach states), thereby compensating costs and, hence, enhancing perceived cognitive efficiency. By contrast, maximal experienced effort arises when there is no detectable progress (i.e., most of the motion of the states lies inside the controllosphere) despite high energy expenditure.

Importantly, task difficulty may impact the ability of the brain to monitor its own activity, eventually shrinking the perceived efficiency of energy resource investments. For example, metacognitive sensitivity decreases when task difficulty increases [125,126]. Thus, allocating processing resources would have a less perceivable effect in a 3-back task than in a 1-back task, which would render 3-back tasks more subjectively effortful.

more unstable states – therefore demanding greater control energy – compared with a 1-back task.

Corresponding and contrasting principles

We have surveyed three distinct theoretical accounts positing how controlled processing gives rise to effort costs: architectural constraints on control processes (bottleneck account), informational changes in the representations of the system regarding instrumental actions (information-theoretic account), and extra energy expenditure required for funneling the system into hard-to-reach network states (network control account). Although clearly distinct in their underlying conceptual bases, there are also notable points of convergence.

For example, under both the bottleneck and network control accounts, effort costs arise from sustaining specific activity patterns (task representations and target states, respectively). Although the network control account provides an explicit, analytic description of the energetic costs of sustaining target states (i.e., it anticipates the experience of costly mental effort based on the input energy needed to stabilize them), the bottleneck account instead assumes that sustaining shared task representations is costly only if it requires prioritizing the use of the underlying resources to the detriment of other alternative tasks that would otherwise bring reward (e.g., mind-wandering; Box 2). Further theoretical and empirical work is needed to assess their overlap. For example, to what extent does the demand for prioritization (of shared task representations) involve difficult-to-reach target brain states?

The link between the information-theoretic and network control accounts is more mathematical and requires the formal reframing of optimal control problems, for example, in terms of variational Bayesian inference [70,71]. In brief, the energetic cost of control inputs can be recast as the KL divergence between the prior and posterior distributions of the equivalent Bayesian inference problem [72]. On this view, gearing the belief of the system toward a posterior state incurs an energetic cost that can be quantified by informational cost. Intuitively, any physical or biological system that implements some form of inference – and thus, actively modifies its internal representations or belief states – will consume energy. Hence, if prior and posterior representations are further apart, more energy is consumed in the process of inference. Therefore, the information-theoretic and network control views are similar in that they both posit a metabolic origin for effort costs. In fact, this is a core assumption of the information-theoretic view [53], which relies on the principles of energy-efficient neural codes [73,74] to justify the informational costs of effort. That the KL divergence between prior and posterior distributions approximates the energy cost of inference is not trivial; in fact, this approximation is compelled to neglect some important aspects of the biology of the system. For example, the cost of modifying the belief state may depend upon the biological implementation of a given belief change, which may be different for different belief types (above and beyond the KL divergence between prior and posterior distributions).

Considering the three views together may also yield useful intuitions. For example, under the information-theoretic account, effort costs arise from updating representations (from priors to posteriors), which may conflict with the notion, under the bottleneck view, that the subjective experience of mental effort originates in sustaining shared representations (as opposed to flexibly updating them). Interestingly, this tension may be resolved by the network control account, which suggests that energy is consumed both in transiently reaching and sustaining target brain states. While the network control account may provide a general framework that could unify these perspectives, it makes no specific commitment to representation of information, or how inferences over representations unfold – which, of course, may shape the information-processing constraints that give rise to effort costs in the first place – but rather, is a general set of principles governing complex systems.

Experimental approaches to adjudicating between views

How might we test the contrasting predictions of the three accounts? Ideally, one would derive quantitative predictions that can be evaluated on common empirical grounds. To this end, we adopt an idealized neuroscientific paradigm, starting with the premise of a 'brain state' defined in terms of brain activity patterns [75] measured using, for example, neuroimaging or electrophysiology. In this view, the transient response of the brain to external or internal events unfolds as a succession of brain states, which, from a dynamical systems perspective, can be represented as a trajectory through a state space defined by (measured) activity across nodes. While general, this definition is sufficient to demonstrate how these three accounts can be empirically tested and compared without committing to specific technique for measuring neural activity. We assume that we can identify individuals' brain states under different *N levels* of the *N*-back task, both before and after stimulus onset, and then consider whether one could use these measures to derive brain-based estimates of an individual's effort costs for each *N* level, under each theoretical account.

To illustrate our reasoning, we rely on simulations of recurrent neural networks (RNNs), trained to perform either the 1-back or the 2-back task ('1-back RNN' and '2-back RNN', respectively) while also processing random distractors. Recall that the *N*-back requires deciding whether the current stimulus matches (a 'target') or differs from (a 'nontarget') the stimulus presented *N* trials earlier, a comparison that can only be made after the current stimulus is presented. In our simulations, we regress trial-by-trial 1-back and 2-back responses concurrently against RNN activity after stimulus onset, defining two response encoding subspaces – projections of RNN activity that maximally discriminate 1-back and 2-back responses, respectively – and four trial types (defined, factorially, by correct 1-back and 2-back responses). We then project the trial-by-trial RNN activity onto these two encoding subspaces before and after stimulus onset, separately for 1-back and 2-back RNNs (Figure 2). Typically, post-stimulus states cluster around specific regions in the encoding space, specific to the history of stimuli (i.e., which item occurred one or two positions back in the sequence) and to their associated response type (i.e., same or different than the current stimulus). As expected, the 2-back RNN exhibited four distinct clusters that separate all trial types (after stimulus onset), whereas the corresponding clusters of the 1-back RNN only separated trials that differed in terms of 1-back responses. As we will see, this simplified setting affords concrete instantiations of the above-described accounts, amenable to empirical neuroscientific testing.

Under the information-theoretic account, effort costs increase with the KL divergence between pre-stimulus and post-stimulus brain states. In our simulations, we observed a larger pre–post KL divergence for the 2-back task compared with the 1-back task. This occurs primarily because the RNN does not need to separate representations for the irrelevant (2-back) stimulus when trained on the 1-back task. In turn, there was no reduction of uncertainty regarding the irrelevant item from pre- to post-stimulus brain states. However, in the 2-back task, the RNN must represent the currently irrelevant stimulus (the 1-back stimulus) so that the network can respond properly when that stimulus later becomes relevant for responding. Thus, KL divergence increases

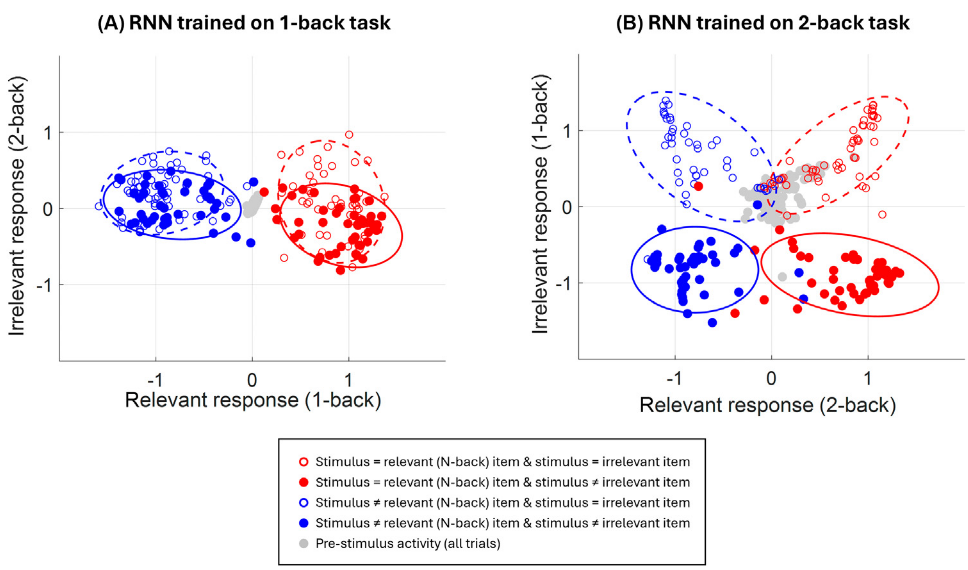

Figure 2. Trial-by-trial recurrent neural network (RNN) states projected on a 2D response encoding subspace; the x-axis represents task-relevant response, whereas the y-axis represents task-irrelevant response. The behavioral response of the RNN to the current stimulus is uniquely determined by its projected state along the task-relevant encoding dimension (e.g., the 1-back response in the case of the 1-back task). This specifies a hyperplane (gray-shaded plane) that separates target and nontarget responses: only state variations that are orthogonal to this frontier can change the response of the network to the stimulus. Gray and white dots depict pre- and post-stimulus RNN activity, respectively. As expected, pre-stimulus activity mostly clusters around the origin (x = y = 0), which indicates that RNNs are maximally uncertain about the upcoming (task-relevant or task-irrelevant) response. In both RNNs, the response to target and nontarget trials cluster into regions that are centered on either side of the frontier (i.e., around x = –1 and x = 1 along the projection onto the task-relevant response dimension). Importantly, the 1-back RNN in (A) does not separate the corresponding clusters along the task-irrelevant response dimension, because the 1-back task does not require RNNs to retain memory traces of 2-back items. In contrast, the 2-back RNN depicted in (B) separates all types of trial along both task-relevant and task-irrelevant dimensions, which means that the post-stimulus responses of the 2-back RNN eventually land into clusters the positions of which strongly depend upon whether the current stimulus is the same as, or different from, either the 1-back or the 2-back item.

with load, as the number of currently irrelevant items on each trial grows with *N,* and this increase in KL divergence should explain increases in subjective effort.

Under the bottleneck account, effort increases with the opportunity cost of prioritizing task-relevant information processing at the expense of alternative activities (e.g., mind-wandering). Crucially, such prioritization is more intensive when the potential for overlapping task representations increases. This implies that opportunity costs increase, particularly when alternative activities must be suppressed due to their disruptive potential. In other words, opportunity costs are high when prioritization demands are high – that is, when competing processes may strongly interfere with one another. In our simulations, such disruptive interference arises when alternative activities (here: processing distractors) induce variations along the task-relevant response dimension. At a first approximation, demand for prioritization can be quantified by the amount of pre-stimulus variability along the task-relevant response dimension, relative to the separation of post-stimulus activity distributions, which provides a functional quantification of bottlenecking in the network. In our simulations, the 2-back RNN exhibited stronger bottlenecking compared with the 1-back RNN because the limited number of network units cannot compress task-relevant information into a small activity subspace that can be shielded from distractors.

Under the network control account, effort costs arise when more energy is needed to guide the system toward post-stimulus brain states that carry task-relevant information. Quantitatively, this depends on brain-wide connectivity, which determines the amount of energy that accumulates along the pre- to post-stimulus trajectory. Qualitatively, however, total control energy will increase with the distance between pre- and post-stimulus activity. In our simulations, the average distance between pre- and post-stimulus clusters was much greater for the 2-back RNN than for the 1-back RNN, owing to the additional pre–post separation that occurs along the irrelevant response dimension in the 2-back, which needs to encode 1-back stimuli for later use. We expect this average distance to increase as a function of *N* level.

Although those scenarios may be simplistic instantiations of the three theories, they suggest a quantitative means by which effort costs can be predicted from measurable brain states. Such brain-derived (‘on-line’) effort costs afford comparison with ‘off-line’ self-reported (subjective) effort ratings (or preferences) to test the ability of each account to predict subjective effort costs.

Concluding remarks

We have considered three recent theories postulating how effort costs arise from distinct foundations, ranging from the opportunity costs implied by the bottlenecking of control representations to the metabolic costs implied by large information shifts needed to perform control-demanding tasks, or achieve rare brain state transitions. The diversity of perspectives suggests that there are multiple reasons why the brain treats cognitive control tasks as costly, but we also identify multiple points of convergence – as indeed not all aspects of these theories are mutually exclusive – and points of disparity (see Outstanding questions). We have also envisioned a neuroscientific approach for testing the various competing predictions of these accounts, which we hope will motivate future empirical work seeking to understand the nature of cognitive effort costs.

Acknowledgments

The authors are grateful to Harrison Ritz, Apoorva Bhandari, Laura Bustamante, and Mickey Inzlicht for helpful comments in the development of this work. This work was supported by NIMH grant R00-MH125021 awarded to A.W.

Declaration of interests

None declared by authors.

Outstanding questions

How might researchers across cognitive psychology and neuroscience reach agreement on a codified operational definition of ‘mental effort’?

How central are opportunity costs to ascribing a source of effort costs – that is, is it possible to define effort costs without appealing to the notion of opportunity costs?

Do each of the three views of the origin of effort costs considered here (information-theoretic, bottleneck, and network control) unequivocally predict that greater working memory demands in classic tasks, such as the *N*-back task, engender increases in subjective (experienced) mental effort?

How might we be able to more ‘objectively’ index an individual’s (on-line) experienced effort costs as they complete a demanding task, both behaviorally and neurally, and what important conceptual methodological and challenges are faced by researchers seeking to do this?

References

1. Allport, A. (1993) Attention and control: have we been asking the wrong questions? A critical review of twenty-five years. In *Attention and Performance XIV: Synergies in Experimental Psychology, Artificial Intelligence, and Cognitive Neuroscience* (Meyer, D.E. and Kornblum, S., eds), pp. 183–218, MIT Press
2. Kahneman, D. (1973) *Attention and Effort*, Prentice-Hall
3. Simon, H.A. (1997) *Models of Bounded Rationality: Empirically Grounded Economic Reason*, MIT Press
4. Gopher, D. and Donchin, E. (1986) Workload: an examination of the concept. In *Handbook of Perception and Human Performance* (2), pp. 1–49
5. Dehaene, S. *et al.* (2017) What is consciousness, and could machines have it? *Science* 358, 486–492
6. Watt, R. (1988) *Visual Processing: Computational Psychophysical and Cognitive Research*, Psychology Press
7. Schneider, W. and Shiffrin, R.M. (1977) Controlled and automatic human information processing: I. Detection, search, and attention. *Psychol. Rev.* 84, 1
8. Kool, W. and Botvinick, M. (2018) Mental labour. *Nat. Hum. Behav.* 2, 899–908
9. Westbrook, A. and Braver, T.S. (2015) Cognitive effort: a neuroeconomic approach. *Cogn. Affect. Behav. Neurosci.* 15, 395–415
10. Ferrero, G. (1894) L'inertie mentale et la loi du moindre effort. *Rev. Philos. Fr. LÉtranger* 37, 169–182
11. Hull, C.L. (1943) *Principles of Behavior: An Introduction to Behavior Theory*, Appleton-Century
12. Samuelson, P.A. (1938) A note on the pure theory of consumer's behaviour. *Economica* 5, 61–71
13. Varian, H.R. (1984) *Microeconomic Analysis*, W.W. Norton
14. Shenhav, A. *et al.* (2017) Toward a rational and mechanistic account of mental effort. *Annu. Rev. Neurosci.* 40, 99–124
15. Devine, S. *et al.* (2023) More than a feeling: physiological measures of affect index the integration of effort costs and rewards during anticipatory effort evaluation. *Cogn. Affect. Behav. Neurosci.* 23, 1129–1140
16. David, L. *et al.* (2024) The unpleasantness of thinking: a meta-analytic review of the association between mental effort and negative affect. *Psychol. Bull.* 150, 1070–1093
17. Inzlicht, M. *et al.* (2018) The effort paradox: effort is both costly and valued. *Trends Cogn. Sci.* 22, 337–349
18. Kraus, J. *et al.* (2025) The pleasure of effort: cognitive challenges trigger hedonic physiological responses. *Ann. N. Y. Acad. Sci.* 1546, 100–111
19. Clay, G. *et al.* (2022) Rewarding cognitive effort increases the intrinsic value of mental labor. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2111785119
20. Mlynski, C. *et al.* (2025) Because it is fun! Individual differences in effort enjoyment belief relate to behavioral and physiological indicators of effort-seeking. *J. Pers. Soc. Psychol.* Published online March 3, 2025. https://doi.org/10.1037/pspa0000443
21. Pütz, S. *et al.* (2024) Physiological predictors of operator performance: the role of mental effort and its link to task performance. *Hum. Factors* 67, 595–615
22. da Silva Castanheira, K. *et al.* (2021) Task-evoked pupillary responses track effort exertion: Evidence from task-switching. *Cogn. Affect. Behav. Neurosci.* 21, 592–606
23. Bermúdez, J.P. and Massin, O. (2023) Efforts and their feelings. *Philos Compass* 18, e12894
24. Halperin, I. and Vigotsky, A.D. (2024) An integrated perspective of effort and perception of effort. *Sports Med.* 54, 2019–2032
25. Bijleveld, E. (2018) The feeling of effort during mental activity. *Conscious. Cogn.* 63, 218–227
26. Braver, T.S. *et al.* (1997) A parametric study of prefrontal cortex involvement in human working memory. *NeuroImage* 5, 49–62
27. Chatham, C.H. *et al.* (2011) From an executive network to executive control: a computational model of the n-back task. *J. Cogn. Neurosci.* 23, 3598–3619
28. Jaeggi, S.M. *et al.* (2010) The concurrent validity of the N-back task as a working memory measure. *Memory* 18, 394–412
29. Master, S.L. *et al.* (2024) Wagers for work: decomposing the costs of cognitive effort. *PLoS Comput. Biol.* 20, e1012060
30. Embrey, J.R. *et al.* (2023) Is all mental effort equal? The role of cognitive demand-type on effort avoidance. *Cognition* 236, 105440
31. Westbrook, A. *et al.* (2013) What is the subjective cost of cognitive effort? Load, trait, and aging effects revealed by economic preference. *PLoS One* 8, e68210
32. Vogel, T.A. *et al.* (2020) Forced choices reveal a trade-off between cognitive effort and physical pain. *eLife* 9, e59410
33. Bustamante, L.A. *et al.* (2023) Effort Foraging Task reveals positive correlation between individual differences in the cost of cognitive and physical effort in humans. *Proc. Natl. Acad. Sci. U. S. A.* 120, e2221510120
34. Bogdanov, M. *et al.* (2025) Noninvasive brain stimulation over the frontopolar cortex promotes willingness to exert cognitive effort in a foraging-like sequential choice task. *J. Neurosci.* 45, e0647242024
35. Cavanagh, J.F. *et al.* (2014) Conflict acts as an implicit cost in reinforcement learning. *Nat. Commun.* 5, 5394
36. Collins, A.G.E. *et al.* (2017) Interactions among working memory, reinforcement learning, and effort in value-based choice: a new paradigm and selective deficits in schizophrenia. *Biol. Psychiatry* 82, 431–439
37. Westbrook, A. *et al.* (2025) Striatal dopamine can enhance both fast working memory, and slow reinforcement learning, while reducing implicit effort cost sensitivity. *Nat. Commun.* 16, 6320
38. Desjardins, S. *et al.* (2025) Context effects in cognitive effort evaluation. *Psychon. Bull. Rev.* 32, 407–416
39. Ashburner, M. and Risko, E.F. (2022) On the influence of evaluation context on judgments of effort. *J. Exp. Psychol. Hum. Percept. Perform.* 48, 790–811
40. Embrey, J.R. *et al.* (2024) Too hard, too easy, or just right? The effects of context on effort and boredom aversion. *Psychon. Bull. Rev.* 31, 2801–2810
41. Wu, R. *et al.* (2022) Do humans prefer cognitive effort over doing nothing? *J. Exp. Psychol. Gen.* 152, 1069–1079
42. Cohen, J.D. *et al.* (1996) A computational approach to prefrontal cortex, cognitive control and schizophrenia: recent developments and current challenges. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 351, 1515–1527
43. Baddeley, A. (1998) The central executive: a concept and some misconceptions. *J. Int. Neuropsychol. Soc.* 4, 523–526
44. Ritz, H. *et al.* (2020) Bridging motor and cognitive control: it's about time! *Trends Cogn. Sci.* 24, 6–8
45. Koch, I. *et al.* (2018) Cognitive structure, flexibility, and plasticity in human multitasking – an integrative review of dual-task and task-switching research. *Psychol. Bull.* 144, 557–583
46. Feng, S.F. *et al.* (2014) Multitasking versus multiplexing: toward a normative account of limitations in the simultaneous execution of control-demanding behaviors. *Cogn. Affect. Behav. Neurosci.* 14, 129–146
47. Musslick, S. and Cohen, J.D. (2021) Rationalizing constraints on the capacity for cognitive control. *Trends Cogn. Sci.* 25, 757–775
48. Musslick, S. and Cohen, J.D. (2019) A mechanistic account of constraints on control-dependent processing: shared representation, conflict and persistence. *Proc. Annu. Meet. Cogn. Sci. Soc.* 41, 849–855
49. Papadopetraki, D. *et al.* (2025) Shielding working memory from distraction is more effortful than flexible updating. *Sci. Rep.* 15, 13512
50. Devine, S. and Otto, A.R. (2022) Information about task progress modulates cognitive demand avoidance. *Cognition* 225, 105107
51. Oberauer, K. (2005) Binding and inhibition in working memory: individual and age differences in short-term recognition. *J. Exp. Psychol. Gen.* 134, 368
52. Oberauer, K. *et al.* (2016) What limits working memory capacity? *Psychol. Bull.* 142, 758
53. Zénon, A. *et al.* (2019) An information-theoretic perspective on the costs of cognition. *Neuropsychologia* 123, 5–18

54. Piray, P. and Daw, N.D. (2021) Linear reinforcement learning in planning, grid fields, and cognitive control. *Nat. Commun.* 12, 4942
55. Melnikoff, D.E. *et al.* (2022) A computational theory of the subjective experience of flow. *Nat. Commun.* 13, 2252
56. Devine, S. *et al.* (2024) Distinguishing between intrinsic and instrumental sources of the value of choice. *Cognition* 245, 105742
57. Douglas, R.J. and Martin, K.A.C. (2004) Neuronal circuits of the neocortex. *Annu. Rev. Neurosci.* 27, 419–451
58. Elston, G.N. (2003) Cortex, cognition and the cell: new insights into the pyramidal neuron and prefrontal function. *Cereb. Cortex* 13, 1124–1138
59. Miller, E.K. and Cohen, J.D. (2001) An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202
60. Passingham, R.E. and Wise, S.P. (2012) *The Neurobiology of the Prefrontal Cortex: Anatomy, Evolution, and the Origin of Insight*, Oxford University Press
61. Betzel, R.F. *et al.* (2016) Optimally controlling the human connectome: the role of network topology. *Sci. Rep.* 6, 30770
62. Gu, S. *et al.* (2015) Controllability of structural brain networks. *Nat. Commun.* 6, 8414
63. Pasqualetti, F. *et al.* (2014) Controllability metrics, limitations and algorithms for complex networks. *IEEE Trans. Control Netw. Syst.* 1, 40–52
64. Bassett, D.S. and Lynall, M.-E. (2013) *Network Methods to Characterize Brain Structure and Function*, WW Norton & Company
65. Parkes, L. *et al.* (2024) A network control theory pipeline for studying the dynamics of the structural connectome. *Nat. Protoc.* 19, 3721–3749
66. Braun, U. *et al.* (2021) Brain network dynamics during working memory are modulated by dopamine and diminished in schizophrenia. *Nat. Commun.* 12, 3478
67. Jarjat, G. *et al.* (2018) What makes working memory traces stable over time? *Ann. N. Y. Acad. Sci.* 1424, 149–160
68. Spaak, E. *et al.* (2017) Stable and dynamic coding for working memory in primate prefrontal cortex. *J. Neurosci.* 37, 6503–6516
69. Wolff, M.J. *et al.* (2020) Drifting codes within a stable coding scheme for working memory. *PLoS Biol.* 18, e3000625
70. Levine, S. (2018) Reinforcement learning and control as probabilistic inference: tutorial and review. *arXiv* Published online May 2, 2018. http://dx.doi.org/10.48550/arXiv.1805.00909
71. Watson, J. (2021) Control as Inference? In *Reinforcement Learning Algorithms: Analysis and Applications* (Belousov, B. *et al.*, eds), pp. 189–206, Springer International Publishing
72. Millidge, B. *et al.* (2020) On the relationship between active inference and control as inference. *Commun. Computer Inf. Sci.* 1326, 3–11
73. Levy, W.B. and Baxter, R.A. (1996) Energy efficient neural codes. *Neural Comput.* 8, 531–543
74. Barlow, H. (1961) Possible principles underlying the transformation of sensory messages. *Sens. Commun.* 13, 217–234
75. Greene, A.S. *et al.* (2023) Why is everyone talking about brain state? *Trends Neurosci.* 46, 508–524
76. Clairis, N. and Pessiglione, M. (2024) Value estimation versus effort mobilization: a general dissociation between ventromedial and dorsomedial prefrontal cortex. *J. Neurosci.* 44, e1176232024
77. Vinckier, F. *et al.* (2022) Elevated effort cost identified by computational modeling as a distinctive feature explaining multiple behaviors in patients with depression. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 7, 1158–1169
78. Frömer, R. *et al.* (2021) Expectations of reward and efficacy guide cognitive control allocation. *Nat. Commun.* 12, 1030
79. Sandra, D.A. and Otto, A.R. (2018) Cognitive capacity limitations and need for cognition differentially predict reward-induced cognitive effort expenditure. *Cognition* 172, 101–106
80. Otto, A.R. and Vassena, E. (2021) It's all relative: reward-induced cognitive control modulation depends on context. *J. Exp. Psychol. Gen.* 150, 306–313
81. Lin, H. *et al.* (2024) An experimental manipulation of the value of effort. *Nat. Hum. Behav.* 8, 988–1000
82. Bogdanov, M. *et al.* (2022) Cognitive effort exertion enhances electrophysiological responses to rewarding outcomes. *Cereb. Cortex* 32, 4255–4270
83. Chong, T.T.-J. *et al.* (2017) Neurocomputational mechanisms underlying subjective valuation of effort costs. *PLoS Biol.* 15, e1002598
84. Le Heron, C. *et al.* (2018) Distinct effects of apathy and dopamine on effort-based decision-making in Parkinson's disease. *Brain* 141, 1455–1469
85. Richter, M. *et al.* (2016) Three decades of research on motivational intensity theory: what we have learned about effort and what we still don't know. In *Advances in Motivation Science* (3) (Elliot, A.J., ed.), pp. 149–186, Elsevier
86. Silvestrini, N. *et al.* (2023) An integrative effort: bridging motivational intensity theory and recent neurocomputational and neuronal models of effort and control allocation. *Psychol. Rev.* 130, 1081–1103
87. Gendolla, G.H.E. *et al.* (2019) Advancing issues in motivation intensity research: updated insights from the cardiovascular system. In *The Oxford Handbook of Human Motivation* (Ryan, R.M., ed.), pp. 373–392, Oxford University Press
88. Gendolla, G.H.E. (2025) The intensity of mental effort: 'the heart does not lie.'. *Synthese* 205, 231
89. Otto, A.R. *et al.* (2022) Is the juice worth the squeeze? Learning the marginal value of mental effort over time. *J. Exp. Psychol. Gen.* 151, 2324–2341
90. Grahek, I. *et al.* (2023) Learning when effort matters: neural dynamics underlying updating and adaptation to changes in performance efficacy. *Cereb. Cortex* 33, 2395–2411
91. Kurzban, R. *et al.* (2013) An opportunity cost model of subjective effort and task performance. *Behav. Brain Sci.* 36, 661–679
92. Schulze, C. *et al.* (2025) A timeline of cognitive costs in decision-making. *Trends Cogn. Sci.* 29, 827–839
93. Lin, H. *et al.* (2022) The average reward rate modulates behavioral and neural indices of effortful control allocation. *J. Cogn. Neurosci.* 34, 2113–2126
94. Otto, A.R. and Daw, N.D. (2019) The opportunity cost of time modulates cognitive effort. *Neuropsychologia* 123, 92–105
95. Agrawal, M. *et al.* (2022) The temporal dynamics of opportunity costs: a normative account of cognitive fatigue and boredom. *Psychol. Rev.* 129, 564–585
96. Dora, J. *et al.* (2022) The effect of opportunity costs on mental fatigue in labor/leisure trade-offs. *J. Exp. Psychol. Gen.* 151, 695–710
97. Badre, D. *et al.* (2021) The dimensionality of neural representations for control. *Curr. Opin. Behav. Sci.* 38, 20–28
98. Gershman, S.J. and Lai, L. (2021) The reward-complexity trade-off in schizophrenia. *Comput. Psychiatry* 5, 38–53
99. Kikumoto, A. *et al.* (2024) A transient high-dimensional geometry affords stable conjunctive subspaces for efficient action selection. *Nat. Commun.* 15, 8513
100. Rigotti, M. *et al.* (2013) The importance of mixed selectivity in complex cognitive tasks. *Nature* 497, 585–590
101. Ritz, H. *et al.* (2022) Cognitive control as a multivariate optimization problem. *J. Cogn. Neurosci.* 34, 569–591
102. Barak, O. *et al.* (2013) The sparseness of mixed selectivity neurons controls the generalization–discrimination trade-off. *J. Neurosci.* 33, 3844–3856
103. Herculano-Houzel, S. (2011) Scaling of brain metabolism with a fixed energy budget per neuron: implications for neuronal activity, plasticity and evolution. *PLoS One* 6, e17514
104. Wang, Z. *et al.* (2010) Specific metabolic rates of major organs and tissues across adulthood: evaluation by mechanistic model of resting energy expenditure1234. *Am. J. Clin. Nutr.* 92, 1369–1377
105. Palombit, A. *et al.* (2022) Variability of regional glucose metabolism and the topology of functional networks in the human brain. *NeuroImage* 257, 119280
106. Attwell, D. and Laughlin, S.B. (2001) An energy budget for signaling in the grey matter of the brain. *J. Cereb. Blood Flow Metab.* 21, 1133–1145
107. Lennie, P. (2003) The cost of cortical computation. *Curr. Biol.* 13, 493–497
108. Jamadar, S.D. *et al.* (2025) The metabolic costs of cognition. *Trends Cogn. Sci.* 29, 541–555

109. Harris, J.J. *et al.* (2012) Synaptic energy use and supply. *Neuron* 75, 762–777
110. Howarth, C. *et al.* (2012) Updated energy budgets for neural computation in the neocortex and cerebellum. *J. Cereb. Blood Flow Metab.* 32, 1222–1232
111. Raichle, M.E. and Gusnard, D.A. (2002) Appraising the brain's energy budget. *Proc. Natl. Acad. Sci. U. S. A.* 99, 10237–10239
112. Padamsey, Z. and Rochefort, N.L. (2023) Paying the brain's energy bill. *Curr. Opin. Neurobiol.* 78, 102668
113. Aiello, L.C. and Wheeler, P. (1995) The expensive-tissue hypothesis: the brain and the digestive system in human and primate evolution. *Curr. Anthropol.* 36, 199–221
114. Mery, F. and Kawecki, T.J. (2005) A cost of long-term memory in *Drosophila*. *Science* 308, 1148
115. Laughlin, S.B. *et al.* (1998) The metabolic cost of neural information. *Nat. Neurosci.* 1, 36–41
116. Sterling, P. and Laughlin, S. (2015) *Principles of Neural Design*, MIT Press
117. Plaçais, P.-Y. and Preat, T. (2013) To favor survival under food shortage, the brain disables costly memory. *Science* 339, 440–442
118. Burgess, C.R. *et al.* (2018) Gating of visual processing by physiological need. *Curr. Opin. Neurobiol.* 49, 16–23
119. Livneh, Y. *et al.* (2017) Homeostatic circuits selectively gate food cue responses in insular cortex. *Nature* 546, 611–616
120. Vishwanath, A.A. *et al.* (2024) Mitochondrial $Ca2^{+}$ efflux controls neuronal metabolism and long-term memory across species. *bioRxiv* Published online February 1, 2024. https://doi.org/10.1101/2024.02.01.578153
121. Csikszentmihalyi, M. (2008) *Flow: The Psychology of Optimal Experience*, Harper Perennial Modern Classics
122. Vuorre, M. and Metcalfe, J. (2016) The relation between the sense of agency and the experience of flow. *Conscious. Cogn.* 43, 133–142
123. Hoffman, B. (2012) Cognitive efficiency: a conceptual and methodological comparison. *Learn. Instr.* 22, 133–144
124. Holroyd, C.B. (2024) The controllosphere: the neural origin of cognitive effort. *Psychol. Rev.* 132, 603–631
125. Fleming, S.M. and Lau, H.C. (2014) How to measure metacognition. *Front. Hum. Neurosci.* 8, 443
126. Norman, E. (2020) Why metacognition is not always helpful. *Front. Psychol.* 11, 1537