

Neural dynamics of cognitive control: Current tensions and future promise

Dale Zhou^{1,2,3}, Danielle Cosme^{4,5,6}, Yoona Kang^{4,7}, Ovidia Stanoi^{4,5,6}, David M. Lydon-Staley^{4,8}, Peter J. Mucha⁹, Emily B. Falk^{4,5,6}, Kevin N. Ochsner¹⁰, and Dani S. Bassett¹¹⁻¹⁷

¹Department of Neurobiology and Behavior, University of California, Irvine

²University of California, Irvine, Center for the Neurobiology of Learning and Memory, Qureshey Research Laboratory, Irvine, 92697, United States

³Department of Cognitive Sciences, University of California, Irvine

⁴Annenberg School for Communication, University of Pennsylvania

⁵Department of Psychology, University of Pennsylvania

⁶Marketing Department, Wharton School, University of Pennsylvania

⁷Rutgers, The State University of New Jersey

⁸Leonard Davis Institute of Health Economics, University of Pennsylvania

⁹Department of Mathematics, Dartmouth College

¹⁰Department of Psychology, Columbia University

¹¹Department of Bioengineering, School of Engineering and Applied Sciences, University of Pennsylvania

¹²Department of Physics & Astronomy, College of Arts and Sciences, University of Pennsylvania

¹³Department of Neurology, Perelman School of Medicine, University of Pennsylvania

¹⁴Department of Electrical & Systems Engineering, School of Engineering and Applied Sciences, University of Pennsylvania

¹⁵Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania

¹⁶Santa Fe Institute

¹⁷To whom correspondence should be addressed: dsb@seas.upenn.edu

Abstract

Cognitive control is a suite of processes that helps individuals pursue goals despite resistance or uncertainty about what to do. Deficits of cognitive control underlie compulsive or risky behavior, as well as other clinical challenges associated with difficulties in regulating impulses, attention, thoughts, and feelings. Although cognitive control has been extensively studied as a dynamic feedback loop of perception, valuation, and action, it remains incompletely understood as a cohesive dynamic and distributed neural process. Here, we critically examine the history of and advances in the study of cognitive control, including how metaphors and cultural norms of power, morality, and rationality are intertwined with definitions of control, to consider holistically how different models explain which brain regions act as controllers. Controllers, the source of top-down signals, are typically localized in regions whose neural activations implement elementary component processes of control, including conflict monitoring and behavioral inhibition. Top-down signals from these regions guide the activation of other task-specific regions, biasing them towards task-specific activity patterns. A relatively new approach, network control theory, has roots in dynamical systems theory and systems engineering. This approach can mathematically show that controllers are regions with strongly nested and recurrent anatomical connectivity that efficiently propagate top-down signals, and precisely estimate the amount, location, and timing of signaling required to bias global activity to task-specific patterns. Importantly, the theory converges with established findings, provides new mathematical tools and intuitions for understanding control loops across levels of analysis, and naturally produces graded predictions of control across brain regions and modules of psychological function that have been unconsidered, marginalized, or indirectly linked. We describe how psychological and network control approaches converge and diverge, noting directions for future integration that could strengthen and sharpen our understanding and predictions of how the brain instantiates cognitive control.

1 Introduction

Meeting the demands of everyday life requires the ability to flexibly pursue all manner of goals, overcoming obstacles and resisting distractions along the way. In psychology, this ability is known as cognitive control, which is as crucial for single-minded striving toward one end as it is for flexibly changing aims [1]. Psychological definitions of control frame it as a suite of processes for prioritizing the pursuit of goals that may conflict with competing thoughts, emotions, impulses, motivations, desires, other goals, and behaviors. For example, cognitive control underlies the ability to inhibit responses that conflict with goals—representations of long-term and momentary outcomes—while also pursuing responses that align with them. To manage such

conflicts, several subprocesses are thought to be involved, including conflict monitoring, cognitive flexibility, task switching, working memory, selective attention, inhibitory control, and learning [2, 3, 4, 5, 6, 7, 8, 9]. Some goals, reached via habitual or automatic behaviors (e.g., thoroughly learned tasks), do not need cognitive control and using it to guide these behaviors would be counterproductive or an inefficient use of effortful control resources [10, 11, 12, 13, 14]. Pursuing some goals rather than others means weighing the value of allocating resources and time. Given that time is finite, as are the resources one can devote to control, this points towards the necessity of computing the value of exerting control [15, 16, 17, 18]. Once seen as a fatigable willpower akin to tired muscles [19], control energy is now understood in terms of allocating mental resources including attention, working memory, motivation, and effort [20].

The importance of cognitive control is evident in its broad conceptual scope and applicability to a wide range of behavioral contexts, from the kitchen and cashier to the classroom, clinic, and courtroom, and even to one's own conscience. Conceptually, cognitive control is intertwined with broader behavioral constructs such as executive function, self-regulation, self-discipline, willpower, volition, delayed gratification, risk-taking, and emotion regulation (**Figure 1**) [21, 22, 23, 11, 24, 25, 26]. Alterations in control function are hallmarks of many neurological and psychiatric conditions, contributing to distress and challenges to daily function related to loss of agency and autonomy [27, 28, 29, 30]. Aspects of cognitive control have been analogized to a “moral muscle” related to a strength of will to resist temptations [31, 32, 33, 34, 35], and research on the developmental trajectory of cognitive control has influenced legal standards for personal responsibility [36, 37, 38]. Finally, control is thought to have a causal power over actions that paves the way for desired outcomes in school, health, finances, and criminal justice with consequences for individuals and societies [39, 40, 41, 42, 22]. These applications highlight a powerful normative worldview about cognitive control historically rooted in beliefs about power, morality, and rationality [19, 43, 33], which require critical reexamination [44, 45].

Although much progress has been made towards characterizing cognitive control at multiple levels of analysis, many open questions remain. At present, working descriptions of cognitive control make different assumptions about what behaviors to focus on, the kinds of processes or representations involved, and how neural substrates instantiate control processes and representations. As such, questions remain about the goals and behaviors that exemplify “good” cognitive control, whether control processes and abilities are domain specific vs. domain general, and whether and how neural systems implement putative processes and mental representations. Varied and disparate working definitions of cognitive control represent major efforts from diverse perspectives to address these and related questions (**Figure 2**) [24, 48, 25, 49, 9, 50].

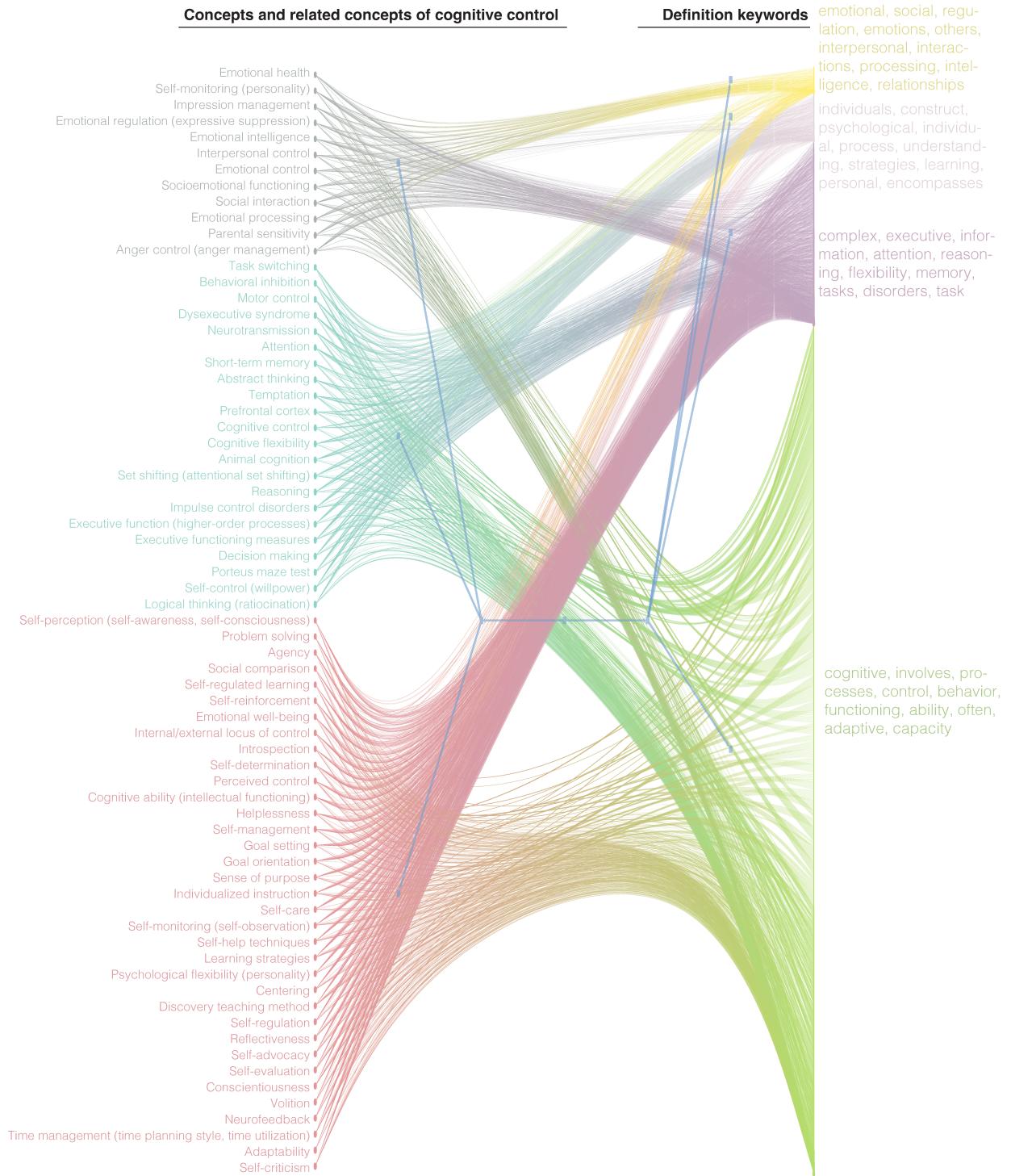


Figure 1: Self-regulation concepts. Conflicting and fragmented working definitions of cognitive control remain a major challenge for developing cohesive theories, with many concepts used interchangeably despite continuous calls for clarity. Inconsistency is unavoidable as a force of scientific progress. In addition to pursuing more ecologically relevant tasks, clearer definitions, and consensus, understanding may come from embracing the plurality of basic scientific and applied goals across disciplinary boundaries. The American Psychological Association Thesaurus of Psychological Index Terms®

Figure 1: provides concepts directly related to the concepts of “cognitive control” and “self-regulation,” as well as the terms’ definitions [46, 24]. Colors indicate groupings of the concept nodes based on community detection of the concept-by-keyword co-occurrence network [47]. Concept nodes are directly or indirectly related to the “cognitive control” and “self-regulation” concepts, such as self-control, executive function, and agency. For the keyword nodes, we show the most frequent words within each group. Edges indicate co-occurrence. The blue tree network indicates hierarchical community structure. See Appendix 4.1 for methodological details.

Although psychologists mostly agree on what kinds of behaviors require cognitive control, conceptualizations of underlying processes still largely involve mapping individual component processes onto individual neural systems. This approach has identified various kinds of putative systems underlying control functions, such as working memory or conflict monitoring. However, this componential approach is difficult to reconcile with recent advances in computational, systems, and network neuroscience that emphasize parallel processing with dynamic and distributed representations. As such, current models of control can seem more akin to a chimeric list of behaviors and component parts rather than a coherent and holistic explanation of how complex behaviors emerge from control processes implemented by interacting neural systems.

By contrast, a neurobiological process account that highlights cognitive control as a distributed and dynamic network function could strengthen cohesion across levels of analysis [51, 52, 53]. A distributed account explicitly models the collective contribution of regions, including those not yet recognized by current paradigms. A dynamic account centers how control operates from moment-to-moment as a function of the evolution of collective activity, rather than only as a function of a terminal state where regions (co-)activate past a threshold. A network account models the brain as a complex system of units (nodes) interconnected by physical pathways (edges), and provides tools to formally quantify collective dynamics across too many regions to grasp without mathematical abstraction. At the intersection of distributed, dynamic, and network models is network control theory.

Similar to other major advances [54], the value of network control theory does not lie primarily in the novelty of its core components—nodes, edges, dynamics, and control targets—but in its explicit formalization of these elements in an interpretable mathematical framework. This endeavor makes it possible to precisely characterize the kinds of control a brain region can exert over its activity toward both known and yet unknown control targets. Network control theory helps synthesize many key ideas in the neuroscience of cognitive control in a precise and mathematically coherent account, linking the processes underlying regulation, development, hierarchy, and cost [55, 56, 57, 58, 59, 60, 61, 62, 63, 54, 64, 65], while producing new predictions for unconsidered control components, processes, and targets. For example, network control theory advances understanding of cognitive control by identifying key brain regions and their development,

predicting when and how much control is required, and extending insights across species. It improves behavioral and psychiatric prediction beyond simpler network metrics (e.g. the strength of a region’s connectivity), and offers a framework for studying and modeling neurological and psychiatric (pharmacological, neurostimulation, psychological) interventions.

Goals of the review

Philosophers and scientists have theorized extensively about how people form goals, make decisions to support them, and flourish through attaining them. What constitutes a good life? Where do goals come from? How do we know if we are achieving them? Defining cognitive control becomes more difficult when both scientific and cultural frameworks constrain how we think about it, historically conflating control with willpower, relying on metaphors about social structures of power, and focusing on individual agency rather than collectives. When theories of cognitive control are shaped by the very social and psychological forces they aim to explain, the risk is that beliefs, values, and assumptions that have become second nature are embedded in our models of nature. How might we move beyond some of these assumptions and expand our ability to see how life actually works and as importantly how it could work? Taking a step back to describe how energy flows in the brain and moves people through decisions and actions with the mathematical language of network control theory offers new insight into how control works and how it might be leveraged to promote flourishing [66, 67, 68].

We first define key ideas across psychological and neuroscientific theories of cognitive control, and then highlight tensions which motivate a new conceptualization of cognitive control. For more detail on core psychological concepts, we point readers to literature on the psychology of goal pursuit and self-regulation [69, 26, 70], computational models of cognitive control [71, 72], syntheses of cognitive control across regions and units of analysis [73, 74, 75, 76], the role of specialized brain regions (e.g., prefrontal cortex [57, 77, 78], dorsal anterior cingulate cortex [54, 79], and insular regions [80, 81]) in allocating control [82], component processes across cognitive control and executive function tasks [5, 83, 11, 8, 6], and the neurodevelopment of cognitive control [84, 85, 86, 61, 63].

The next section conveys mathematical intuitions underlying network control theory, showing what is unique about the formulation with no assumption of prerequisite mathematical background. For more extensive mathematical treatments of the topic, see literatures on the general theory and applications [87, 88, 89], applications to neuroscience [90, 67, 91], didactic textbook chapters [92, 93], practical tutorials and code [94, 95], application to animal models including directed connectomes of *C. elegans* and mice [96, 95],

and applications to perturbing and predicting neural dynamics [97, 98, 99]. In the final section, we outline how each approach contributes to the other and potential future efforts to interdigitate the two constructs. When we refer to the psychological perspective of control, we use the term “cognitive control,” whereas we refer to the network science perspective, we use the term “network control.” When referring to the broader concept of self-regulation, we simply say “control.”

2 Cognitive Control: Constructs and Correspondence

The psychological construct of cognitive control

“If in general we class all springs of action as propensities on the one hand and ideals on the other, the sensualist never says of his behavior that it results from a victory over his ideals, but the moralist always speaks of his as a victory over his propensities. [...] And if a brief definition of ideal or moral action were required, none could be given which would better fit the appearances than this: It is action in the line of the greatest resistance.”

— Ref. [100]

Historical origins

Why would anyone act contrary to their own goals and ideals [101]? It has long been recognized that achieving feats of self-control is difficult because different mental, emotional, and motivational processes conflict with one another. Plato conceptualized the mind as a team of horses representing passion and appetite controlled by a charioteer representing reason [102]. Without the charioteer, the horses run wild and the mind cannot pursue virtuous purposes. Continuing in this contest between “dual processes,” the study of cognitive control has historically been framed as a conflict between uncontrolled and controlled aspects of the mind: automatic vs. deliberate; hot vs. cold; reactionary vs. willful; stimulus-driven vs. goal-driven; impulsive vs. inhibited; id vs. ego; distracted vs. attentive; off-task vs. on-task, and habitual vs. effortful. Such contests are frequently moralized as conflict between inner demons versus better angels, wherein the victory of virtues over vices drives human progress [103, 104, 105, 106, 33].

Over a century ago, at the dawn of psychology, control was thought to tip the scales in these conflicts. This idea was captured by a simple yet enduring formula: cognitive control energy or effort E decides the balance between habitual instincts, impulses, motives, pleasures, and pains as propensities P and our ideals I [100]. E is needed when $I < P$; that is, when habits override ideal goals. Cognitive control allows for

$$I + E > P.$$

Recent advances

Although oversimplified, this basic formulation has been expanded or rearranged to explain many aspects of control behavior. E is not a unitary resource of will but a decomposable process with multiple elementary components [107, 108]. E is partly about weighing the value of alternatives and monitoring progress [109, 110, 111, 71, 112, 72, 113, 114], a motivated and context-sensitive valuation of how to allocate time and let go of lost opportunities. This does not merely rebrand willpower as willingness to apply effort; this E is a subjective valuation that weighs the costs and benefits of I as contexts change, tracking when I promises greater reward, I is more probable, and reduced E can lead to better control outcomes [115, 116, 14]. E can also act as a divisor to P , rearranging the inequality to bring E to the other side as an inhibitory gate so that $I > \frac{P}{E}$ [41, 4, 5, 54, 48, 50]. Other components amplify or diminish E , such as emotions and motivations, and P is a function of many states and traits, such as risk taking, sensation seeking, and conscientiousness [117, 118, 119, 54, 120, 9, 121]. Taken together, E is not a unitary input to tip the balance between (or among) habits and goals, but is a series of calculated choices shaped by separate but interacting component processes [122].

Component processes

Varied ideas have arisen across fields about the nature of these component processes (**Figures 2 and 3**). Many psychologists view control as emerging from inhibition of automatic or goal-irrelevant responses, updating and monitoring working memory, and shifting attention between tasks [5, 11, 50]. Cognitive psychologists view control processes as a function of learning to make reward-based decisions by applying selective attention and working memory to monitor and gate ongoing action when there is uncertainty about what to do until sustained or momentary goals are completed [109, 123, 110, 124, 71, 125, 126, 127, 72]. Personality and social psychologists view control as the ability to regulate driving and restraining forces, supported by certain traits and feelings of autonomy, competence, and relatedness, in service of a rich hierarchy of personal goals of differing abstractness, generality, and duration [128, 129, 32, 21, 130, 131]. Expectancy-value theorists in motivation psychology view control as driven by goals that have sufficient value balanced with confidence about the self-efficacy towards expected attainment in service to a system of needs and drives [132, 133, 134, 135, 136, 106, 130]. Neuroeconomists view control processes as a cost-benefit analysis in service of choice that maximizes benefits under costs of alternative opportunities, time,

and mental resources [137, 138, 116, 112, 20, 113].

Shared computations of dynamic feedback loops

Despite these differences, all consider the core control process as a dynamic feedback loop of perception, valuation, and action. This feedback loop describes cognitive control as a process balancing the choice to apply control effort with the value and efficacy of that effort, given the perception of the current situation. Choices influence what new sensory inputs, task stimuli, and options one attends to, and the “perception-action cycle” continues [140, 141, 126, 142, 143]. Over time, effective control is learning the range of actions that reduce the difference, or increase the congruency, between the expected value of choices and the perceived state of affairs [144, 1, 145, 14, 146].

Control theory provides systems engineering concepts to understand how dynamic feedback loops operate at multiple scales of time and abstractness for behavioral goals or targets [129, 147]. Negative feedback reduces the discrepancy between the perceived situation and current goals in order to maintain goal pursuit, while positive feedback amplifies discrepancies to repel further from anti-goals. For instance, thermostats adjust temperature to close the gap with a reference point with a negative feedback loop. Blood clotting involves platelets releasing signals that attract more platelets to the injury site, compounding quickly to create distance from dangerously low levels of blood in a positive feedback loop.

Many aspects of everyday control behavior involve approaching a target goal or avoiding an anti-goal, in accord with negative and positive feedback loops [1]. For example, thinking outside the box involves approaching a creative idea and avoiding familiar and rigid mental sets. Indeed, cognitive computational neuroscience models using supervised learning instantiate positive and negative feedback loops to minimize or maximize multiple cost functions for cognitive control tasks [109, 148, 123, 149, 116, 150]. An effective control system is one that allows for reaching the widest range of routes to goal states efficiently, either in the world or our mind, despite great conflict and given diverse initial contexts and situations. These functional and algorithmic processes guide the investigation of the neural implementation of control.

Neural bases of cognitive control processes

“*Nothing in excess.*”

— Delphic maxim (circa 6th century BCE)

A century-old analogy for the neural bases of cognitive control considers how control regions of the

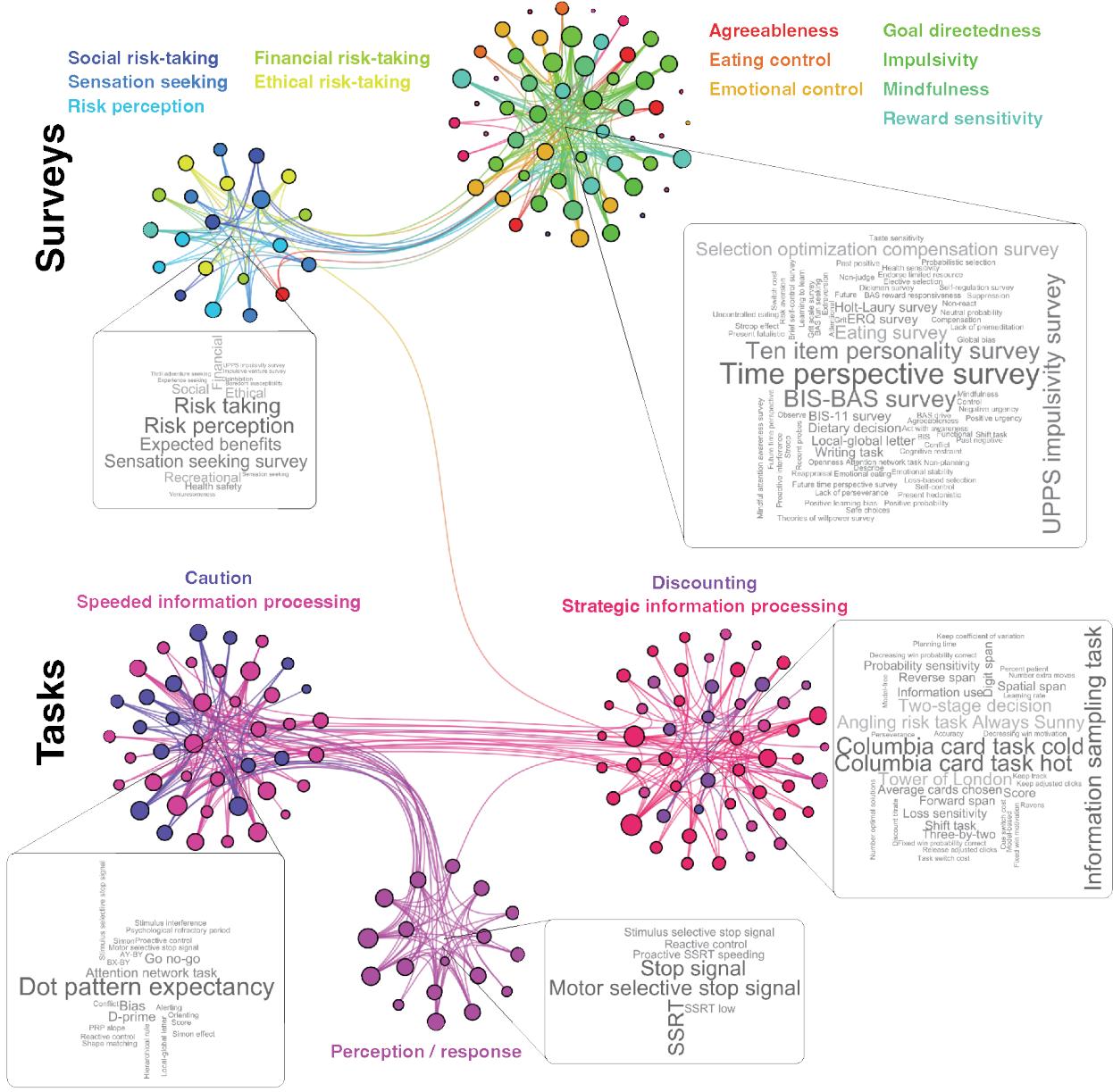


Figure 2: Self-regulation measurements from tasks and surveys. Self-regulation often refers to processes that support goal pursuit broadly, and may benefit from but do not necessarily require certain forms of cognitive control [23]. Current survey and task measures, including cognitive control constructs, exhibit limited coherence (weak connections between task and survey modalities and within conceptual domains). A “cognitive ontology” is visualized using 129 task and 64 survey dependent variables reproduced from previously published data [9]. Nodes are defined as dependent variables which are colored and labeled by 5 task factors and 12 survey factors identified by the published exploratory factor analysis. Edges are defined as partial correlations between behavioral performance or self-report measures from 560 participants across all pairs of dependent variables. Edges are sparse due to moderate regularization ($\alpha = 0.15$) using Graphical Lasso and a partial correlation threshold of $r \geq 0.05$. Nodes are clustered based on a stochastic block model [139]. Word clouds depict the node identities per cluster. See Appendix 4.2 and Ref. [9] for details.

brain act as a way station or switch that changes the flow of train traffic across a system of railroads [100]. Some of the most influential current conceptualizations of control—the biased competition and guided activation frameworks—apply this analogy [57]. The prefrontal cortex can be seen as a central station that implements cognitive control by redirecting trains carrying information about goals to manipulate the flow of command or stop signals at other stations, thereby biasing their activity to produce actions in line with goals [56, 57]. Control signals coordinate traffic by adding trains, scheduling delays and slowdowns, and selectively admitting or banning only certain types of passengers using selective attention, inhibitory, and learning processes. All these activities help to maximize efficiency and minimize interference, redundancy, and conflict by avoiding traffic congestion as much as possible. There is correlative, causal, computational, and neurodevelopmental support for this framework [151, 84, 152, 153, 154, 61, 62, 59, 155].

Monitoring the conflict and congruency of goals and choices to switch tasks

Conflict arises from traffic across the brain, or the interference between simultaneously co-activating or spatially overlapping patterns of activity which represent competing or incongruent choices and goals [2, 156, 114, 145, 157, 158]. Traffic makes cognitive control aversive, cost time, feel difficult, and occupy mental resources [116, 159]. The dorsal anterior cingulate and medial prefrontal cortices detect conflict to signal error and to prioritize efforts, sending conflict signals that stimulate resolution among signals with more control [148, 160, 16]. These efforts are coordinated by goal representations in prefrontal regions, acting as attentional templates to filter or bias downstream neural processing for multiple task rules, goals, opportunity costs, expected values, and levels of effort [161, 162, 163, 164, 165, 57, 58, 166, 167, 54, 168, 169, 170, 171, 172, 173]. For example, dorsolateral prefrontal cortex is thought to bias attention and working memory to promote responses that are appropriate for goals while limiting responses that are inappropriate [174, 57, 77]. Ventromedial prefrontal cortex is thought to predict outcomes relative to goals [175, 176, 57, 54, 8], regulating arousal and affect [177] as well as adapting to new information that is incongruent with previously learned information [178, 179]. Ventrolateral prefrontal cortex is associated with the selection of goal congruent response options among competing alternatives [180, 181]. These specialized prefrontal regions communicate with a broader frontoparietal and cingulo-opercular network [182, 183, 59, 68], including the anterior insula/operculum, dorsal anterior cingulate cortex, and thalamus [184, 185, 186, 187]. Such highly connected regions have many railroad connections, some which make U-turns back to the origin to allow signals to recur stably, that can help to flexibly enact control to modulate conflict, maintain stable control representations, and prioritize activity that is salient or relevant to goal representations [123, 183, 126,

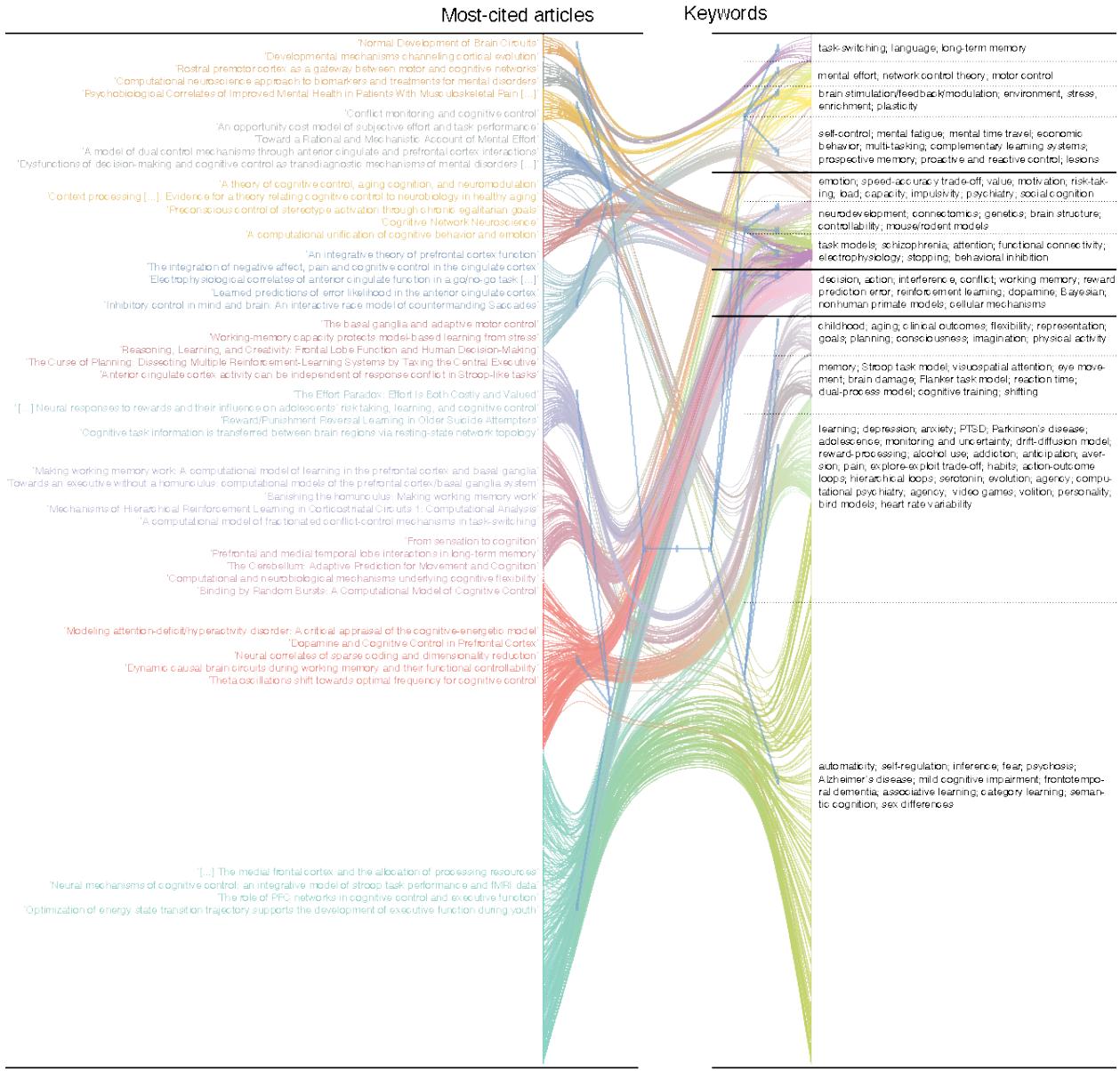


Figure 3: A plurality of concepts from highly cited articles on cognitive control computations. The challenge is to develop a cohesive yet precise framework that can speak to a plurality of concepts at multiple levels of analysis. Visualization of a hierarchy of concepts from 308 highly cited papers that included the keywords “cognitive control” and “computational” from a Web of Science search. Topics determined by a hierarchical stochastic block model of an article-by-keyword co-occurrence network [47]. For article nodes, we show five or fewer randomly selected articles in the community. For the keyword nodes, we show the most frequent words. Edges indicate co-occurrence. The blue tree network indicates hierarchical community structure, where the highest hierarchical level is demarcated by the solid and bold black lines, the middle level is demarcated by solid black lines, and the most granular level is demarcated by dotted black lines. See Appendix 4.3 for details.

59]. As a form of reactive control, competing signals stimulate control signals to resolve conflict, increase resource efficiency, and improve performance by separating previously overlapping representations [157, 125, 6]. Conflict can also be preemptively avoided by proactively integrating multiple neural representations of the control task and goal [173, 188, 189].

Stop signaling and behavioral inhibition

To alter control and conflict in accordance with goals, some inputs and outputs can be suppressed using the inhibitory signaling of cortico-striatal loops and neuromodulatory pathways [190]. Cortico-striatal loops send stop signals from the striatum to cortical regions via thalamic anatomical projections, such as those maintaining information in working memory about goals and the task at hand [191, 192, 193, 194, 195, 196, 4, 197]. In concert with the inferior frontal gyrus (which includes ventrolateral prefrontal cortex), this fronto-striatal stop circuit inhibits automatized behavior to select actions in pursuit of a goal [198, 199, 200, 197]. As such, striatal involvement provides a bottom-up control signal that complicates predominantly top-down views of cognitive control.

Value and salience shapes goal-oriented perception, choice, and learning

Valuation and salience processes also help alter control and conflict activity in accordance with goal representations, calculating which actions to select or suppress in relation to a goal using parallel cortico-striatal loops and neuromodulatory signaling [110, 201, 202, 203, 126, 64, 143, 146]. The ventromedial prefrontal cortex and striatum interact with the dorsal anterior cingulate cortex to support the valuation of the potential courses of action that integrate affective and salience inputs through connections with the anterior insula [201, 80, 54, 141, 186, 187]. At the neuronal level, incoming and outgoing connections are prone to change such that useful transit lines are more heavily traveled and reinforced by positive feedback or are quieted and closed by negative feedback from other stations, a learning-dependent plasticity that helps save, monitor, and improve progress [204, 55, 205, 149, 206].

Brain-wide top-down and bottom-up signaling and neural signatures of cognitive control

Theories place differing importance on top-down signaling relative to other forms of signaling [57, 51] (**Figure 4**). Signals of goals, conflict, and stopping are types of top-down signals that filter, amplify, trigger, and redirect certain bottom-up inputs via communication and control [91]. For instance, bottom-up perceptual signals are biased to select or suppress the parts that are most or least relevant to current goals [56, 207, 208].

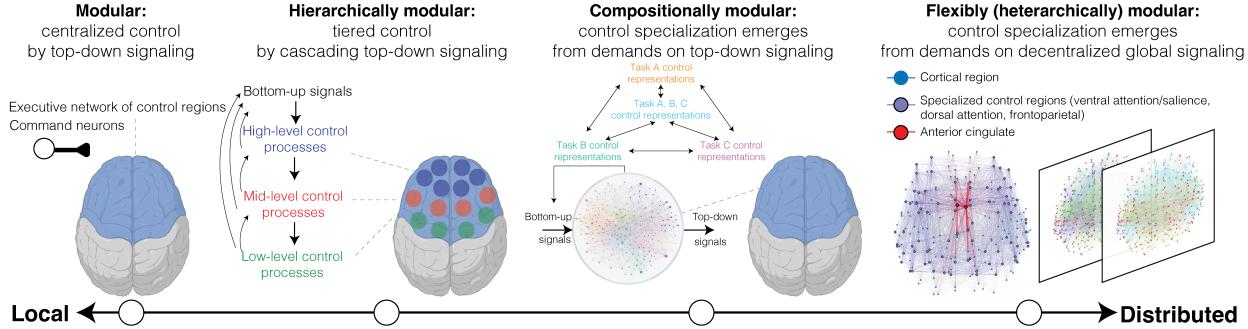


Figure 4: Local and distributed control signaling. Hypothesized control processes place differing importance on centralized versus decentralized signaling [51]. *Modular control* is largely encapsulated and centralized within a few regions with distinct biological composition to output top-down signals. *Hierarchically modular control* partitions processes along a stratified ladder, maintaining separation of function as top-down signals cascade down the rungs and feed back signals to monitor progress. *Compositionally modular control* emerges from interactively integrable representations learned to meet the demands of control behavior to output top-down and feedback signals. *Flexibly modular control* emerges from representations meeting the demands of control behavior engaged by global top-down, bottom-up, lateral, and recurrent signaling. In contrast to a hierarchy, a heterarchy involves varied modules that consolidate and dissipate over time according to demands on global signaling. Control signals evolve simultaneously and in parallel across neural systems. Which signals lead or follow is relative to the perspective of the scientist (analogous to opening a chain loop to define which link leads and which follows; the break sets which element appears first). These hypotheses can be complementary and need not be mutually exclusive—a heterarchy can involve a temporary hierarchy containing compositional representations in hierarchically nested regions with specialized local biological compositions. Brain image is from BioRender.

The collective activity of top-down signals may be enacted by alpha, beta, and theta oscillations from deeper cortical layers, traveling from anterior to posterior regions of the brain [209, 154]. These top-down signals may modulate bottom-up gamma bursts from superficial cortical layers to bias perceptual signals as they propagate to task functional regions [210, 211]. Striatal, insular, and thalamic loops further suggest that motivation, social cognition, reward, and emotions are core signals that dynamically interact with top-down signals, complicating notions of a predominantly top-down hierarchy of control [128, 212, 52, 213, 214, 215]. Recent approaches focusing on neural signatures remain more agnostic to such theories and models of neural systems, and instead classify and decode the patterns of activity across the brain that indicates when a person is actively engaging in cognitive control processes [216, 217, 218].

Summary

Control emerges from top-down selective attention or inhibitory gating processes that bias the inflow and outflow of bottom-up signals to filter, amplify, trigger, or redirect parts of the signal related to goal representations. Continuously monitoring the situation, other regions detect conflict between top-down and bottom-up signals to stimulate more control adjustments according to value, cost, and perceived progress.

Limitations and tensions of current psychological and neuroscientific approaches

“[T]here occurs a kind of progression in industrial discipline moving from paternalistic controls to assembly line, machine paced routines and, finally, to bureaucratically imposed discipline. What is involved is a shift from heteronomous paternalist controls to autonomous, internalized discipline, and identification with corporate goals and values.”

— Ref. [219])

Current constructs have several limitations and tensions at the neural and behavioral levels, as well as at the cultural and normative levels.

Neural and behavioral tensions

First, it remains unclear exactly how the implicated neural activity represents goals, rules, conflict, costs, benefits, options, actions, and feedback, let alone how different perceptual, affective, motivational, motor, and control representations are integrated with them. In computational models that have incorporated neurophysiological processes and aim to model a cognitive control region’s function (e.g. prefrontal cortex), the propagation and transformation of activity allows it to generate cognitive control behavior, dependent on that activity originating from a source that already encodes these kinds of relevant information [55, 220, 148, 221, 222, 149, 14, 223]. Ongoing work aims to explain how cognitive control representations are neurally encoded [193, 224, 225, 211, 189, 157, 173, 188, 226, 227].

Second, what counts as “good” cognitive control is ambiguous at the neural and behavioral levels because it depends on context. Different instances of control behavior can be shown to depend on different neural systems. For example, a cognitive control task that focuses on personally or socially relevant goals elicits default-mode activity compared to tasks that focus on more contrived goals [228, 54]. Moreover, different valuation processes can change what counts as success vs. failure for the exact same control behavior [111, 229, 230]. For example, whether a measure like reaction time, variability, high cost, learning rate, momentary hedonism, or sensation-seeking indicates good or bad control is context-dependent [111, 229, 9]. This ambiguity arises because many models are underconstrained [231], allowing many (sometimes contradictory) ways to explain the same behavior. Fortunately models can incorporate a broader scope of relevant conceptual and neuroscientific variables [111, 229, 113, 232, 14, 233, 82, 234].

Cultural and normative tensions

Third, a limited set of behavioral instances of control tend to be studied, and those instances seems to cater to societal norms and beliefs about control that are desirable or virtuous. Blending the biology and culture of control blurs the line between self-discipline by internally generated goals and discipline imposed by societal norms [219, 235, 236]. When construed as encompassing all goal-directed behavior in general, cognitive control can be used to pursue any goal. Yet out of all the countless possible goals that could be pursued, a relatively small set of norms drive outsized attention to goals that seem selected for fitting a certain set of sociocultural expectations about asceticism, mortification (subduing one’s “earthly nature” and sinful impulses through self-denial), and transgression that can be traced to philosophical and religious literature over two millennia. Striving for temperance, personal responsibility, and meritorious achievement certainly may lead to healthier, happier, and more productive lives and communities [237, 42, 39]. It is perhaps unsurprising that endorsement of such norms is related to normative metrics of success that are partly defined by self-surveillance according to those norms [9]. But are those goals not more properly value judgments, spiritual exhortations, or moral presuppositions instead of a fundamental neurobiological and cognitive function, even though much could be said in their favor [39, 40, 42, 22, 237]? One wonders how experimental tasks, extant control representations, and outcome metrics of success in life would change if creativity, compassion, cooperation, or altruism were elevated. Just as one could expand the scope of goals and flourishing, conversely, one can readily construe goals wherein strong control leads to guilt, shame, isolation, oppression, deceit, and violence. Or one could construe of goals that appear frivolous, fanciful, or impossible [238]. Normative assumptions reflect the current zeitgeist, but such assumptions not only threaten generalizability across studies and cultures [239], but generalizability across time and to other species. There is an ambivalence to cognitive control, more like perception or movement than a virtuous vehicle to prosperity.

Fourth, cognitive control is not a universal recipe to take action and reap rewards. General applicability to all goals tempts a perennially enticing belief in an inner fuel—like willpower—that can get us to wherever we set our minds to: “where there is a will, there is a way” [240]. In contrast, there is inconsistent evidence for broad applicability of control leading to desired outcomes [40, 42, 21, 24, 241, 9, 242], evidence for more specific and local transfer across goals based on the type of control strategy and task [123, 243, 244, 245, 14], and a lack of relationship between control components and putative domain general constructs of intelligence [246]. Willpower is thought to enable mastery over oneself and, with some luck, imposition of one’s will on the environment and others. This is power misidentified as willpower. The temptation to

confuse power for willpower may be explained by how it empowers people to take responsibility for their fate, a link that justifies why certain people deserve power and others do not, to appear greater than their fortune. One problem that this confusion creates for the science of control is constraining the space of hypotheses to indulge what some may already want to believe—that we can achieve whatever we set our minds to, that we can bend the world to our will, and that clear ambition and focused effort fan fortune towards these ends. These are beliefs about power. When applied to cognitive control, those beliefs support the idea that individual differences in control and their purported outcomes are not only justified but biological. Such beliefs in willpower are outmoded in cognitive control [247] and supplanted by dynamic feedback loops of perception, valuation, and action. We do not have worse cognitive control because we lack the strength or capability of will. Rather, worse cognitive control occurs when there is a large discrepancy between what we perceive a situation demands of us and what we have learned to subjectively evaluate as being worth doing, too costly to change, and as opportunities too regrettable to lose given limited time and attention. Nevertheless, a remarkable causal and explanatory power is still often assumed of cognitive control as a solution for the many outcomes that life seems to demand of us, whether or not they arise from our own values. Confidence in this assumption appears disproportionate relative to the construct’s modest internal coherence, its modest predictiveness of outcomes with many possible unconsidered causes, and the negative outcomes of its overuse [248, 9, 239, 14].

Finally, many current conceptualizations of control have an issue known as the “homunculus” problem, because they postulate internal representations of goals and processes that themselves manage control in the same way a person might. While a homunculus is no longer as explicit as it once was—embodied, for example by Plato’s notion of a charioteer of reason or a more recent notion of a central executive CEO who organizes information in their prefrontal office to pursue goals [102, 249]—there is still a passing resemblance to human-like controllers in many models, such as an economist evaluating information with rational cost-benefit analysis to maximize the utility of exerting control. Although these metaphorical sources of control are not explicitly posited by any modern theory [225, 209, 140, 141, 1], they are often implied by or necessary to models’ formulations, from will to utils. This has been the case since the classic model of dynamic test-operate-test-exit feedback loops, which resembles control by business managers. In a dynamic process with top-down authority and hierarchical communication channels, executive goals set worker objectives, progress is monitored and evaluated, and rewards or punishments are delivered [109, 250, 251]. Quantifiable goals and task delegation to specialists maximize the efficient distribution of labor and minimize conflicts that prevent resources from being smoothly allocated, strategies with origins in business management [252, 253, 251].

These are useful and ubiquitous cognitive and cultural tools deserving of application and study, but treating life as if it is a business risks mistaking management metaphors for the nature of self-management. Using metaphors is not necessarily problematic. Metaphors help generate ideas, and they are not necessarily even wrong. After all, humans behave in a world governed by values, traditions, and economics. The issue is when parts of metaphors at the behavioral and psychological level also map out the neural and physical levels. The problem is not that scientists seriously endorse homunculi, but that large knowledge gaps left unexplained about the neural representations of control concepts can implicitly rely on the intuition of homunculi to hammer out and haunt those very voids.

Summary

We often impose cultural ideas of what control should do onto our models of how control works. The CEO is one such idea, traditions that emphasize rational and virtuous control is another example, and the simplistic idea that people succeed in life because of willpower, where willpower is the currency of control, is another example. The field has already reckoned with many of these nuances, but they have a stubborn influence. It is difficult to cleave ideas that have endured for over two millennia. A neurobiological process model of cognitive control can offer mathematical explanations that help address or circumvent some of these challenges [51, 52, 53, 254].

The network control theory construct of control

“It is not necessary, however, to organize one’s Plans in terms of frozen and brittle terminal states. Unlike good problems in science or mathematics, successful living is not a “well-defined problem,” and attempts to convert it into a well-defined problem by selecting explicit goals and subgoals can be an empty deception. [It] is better to plan toward a kind of continual “becoming” than toward a final goal.”

—— Ref. [109]

Originally formalized in the field of systems engineering, control theory has been applied to regulate the behavior of systems in technology, communication, transportation, and robotics [255, 88, 93]. In neuroscience and psychology, different forms of control theory have helped to understand perception, movement, reward-guided decision making, and self-regulation [256, 257, 54, 258, 98, 150, 1]. Here we describe how network control theory can be applied to understand the neuroscience of cognitive control, at a high level with no assumptions of prior mathematical background.

Basics ingredients of network control theory

As noted above, current models of control in psychology and neuroscience (e.g. the biased competition and guided activation models) describe control regions as implementing top-down signals that have a biasing effect on other regions, guiding their activity as if providing instructions for how to move trains in a system of railroads [100, 57, 59]. By contrast, network control theory formalizes these ideas into a cohesive framework by explicitly specifying the kind, amount, location, and timing of activity across the brain expected to enact such control. On this view, an effective control system allows for efficiently reaching the widest range of routes to goal states, given diverse initial contexts and situations. The extant literature on network control theory in neuroscience primarily operates at the level of how control is implemented to efficiently reach the widest range of routes to brain states.

To illustrate how network control theory works, elements of the railroad analogy [100, 57] remain useful, including the railroads, trains, routes, ongoing traffic, target traffic pattern, and train conductor. In place of railroads tracks, neural signals have the hard constraint of continuously propagating across the anatomical white matter pathways between regions, which can be measured by diffusion imaging (**Figure 5A**). Methods to estimate effective connectivity from activity have also been used to model this constraint [259, 260]. These “tracks” converge at different “stations,” or brain regions.

In place of informational “trains,” the *spontaneous dynamics* of each region is determined by propagating threads of activity—signals sent by one region to neighboring regions and still others across their anatomical connections (**Figure 5B**). The magnitude of activity is amplified or constrained according to the strength and volume of the connections it travels across. When multiple threads of activity intersect at the same region, their activity is summed, further propagated to connected neighbors, and decays over time.

Due to the hierarchical hub-and-spoke connectivity of the brain [265], what started as hundreds of threads of activity splits off into hundreds of thousands of simultaneous threads as each thread splits off at more and more regional intersections. Due to the small-world modularity of brain structure [266], after crossing just four or five intersections, threads provide multi-step functional pathways between all regions, although longer functional pathways are far less reliable and more costly to traverse than shorter pathways [262, 263]. The mathematics used to calculate the predicted future behavior of these dynamics precisely formalizes how much, where, and when neural signals have “downstream influence” (**Figure 5C**).

The *control energy* is the amount of additional activity that is needed to account for the difference between the future activity predicted in spontaneous dynamics and the amount of future activity that serves as a *control target*. When network control theory is applied to behavioral control tasks, control targets are

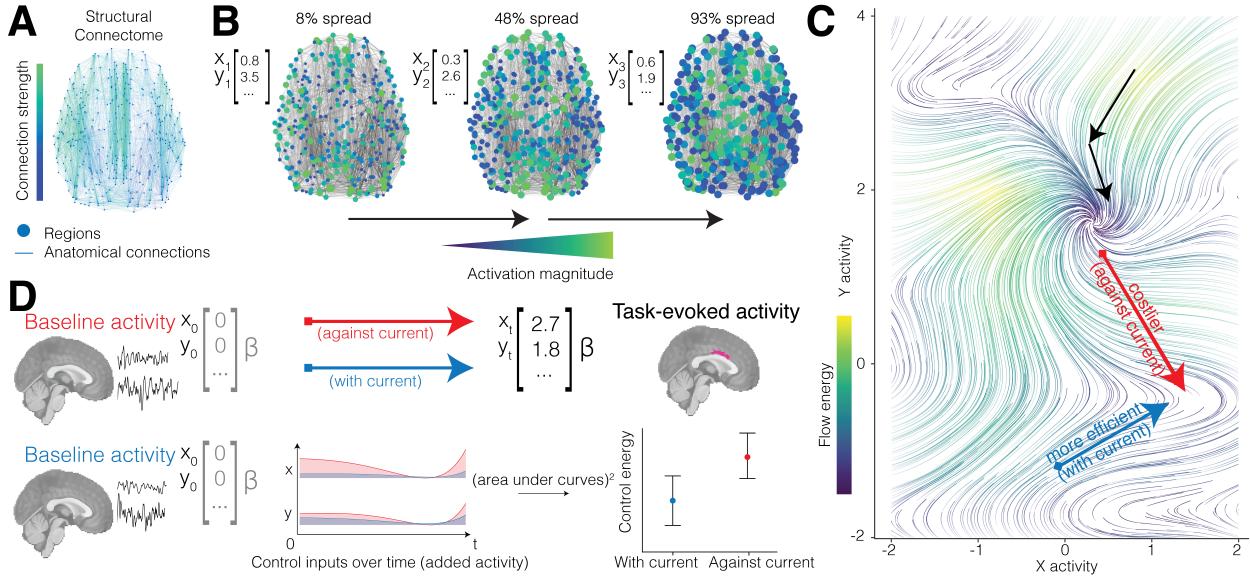


Figure 5: Guided activation according to network control theory models of spontaneous and controlled activity flow.

(A) The anatomical connectivity of the brain, as defined by a structural connectome. **(B)** Neural activity spreads across anatomical connections over time [261, 262, 263]. For illustration, consider just two regions (though all are modeled). From an initial state, the change in activity for region X and region Y over two transitions: (x_1, y_1) to (x_2, y_2) and (x_2, y_2) to (x_3, y_3) . The activity values can be predicted linearly, meaning that regional activity is calculated by taking the source activity multiplied by the connection strength and adding it to the target's activity. Linear dynamical systems theory allows us to solve this calculation across all regions and connections simultaneously and understand how global activity patterns emerge from local regional connectivity profiles per time step. **(C)** A conceptual schematic showing how region X and region Y 's activity flows over time. The two transitions from Panel B are drawn with black arrows. Simulating the predicted activity from many different initial points can provides a comprehensive understanding of the dynamical flow of activity, visualized as streamlines. Streamlines drive activity flow away from zones of higher energy (yellow) to zones of lower energy (blue). The controlled transition between some states (red arrow) must overcome a stronger current of natural dynamics than other states (blue arrow). **(D)** Consider a control problem that starts at two different baselines for the same person. The problem of control is to solve for the minimal time-varying input to add to the natural dynamics to reach the pattern of activity serving as the control target. Here, consider the control target as dorsal anterior cingulate cortex activity evoked during a control behavior that is thought to signal the levels of cognitive conflict and goal congruency. The control energy is the square of the integrated input added over time such that the energetic cost rises steeply per unit of input. Greater energy is needed for the red trajectory because it must overcome stronger currents of natural activity flow than the blue trajectory. Network control theory provides closed form solutions and tractable numerical simulation tools to find optimal solutions for this problem given any selection of initial states and control targets, and does so simultaneously for any set of regions and connections per time point. Figures reproduced and adapted from [262] and [264].

patterns of activity elicited by a goal or cognitive control task, analogous to prior cognitive and computational neuroscience approaches (**Figure 5D**). Defining control targets as future patterns of task-evoked activity can map network control to previously studied activity implicated in the component processes of a cognitive control task (in a standard fMRI analysis, this is a general linear model’s beta coefficient). There are also other sensible choices for defining control targets, including measurement units from machine learning classification or decoding methods that index neural signatures of cognitive control, as well as the evolution of neural activity measured with electrophysiology, spike recording, and calcium imaging [98]. The definition of control target is flexible. Modelers can build in but are not bounded by what is considered as a “correct” behavior in a well-defined task, providing a useful theoretical foundation for yet unconsidered control targets.

The amount of energy needed to attain a target pattern of activity can be likened to a calculation of the energy needed to redirect a sailboat towards a target direction by reorienting its sail against the naturally changing currents of wind and water (**Figure 5C-D**). The “energy” in control energy refers to the cost of additional activity ramping up exponentially. This makes modeled activity evolve over time within relatively stable bounds and prohibits exceptionally large deviations which are not empirically observed in biology due to biophysical constraints [267]. Despite the different name, energy has the same behavior as the propagating threads of activity described above. Different regions have different available solutions to the control problem of how much, where, and when energy input would usefully add to spontaneous activity to control the evolution of activity to control targets with the allocation of as little control energy as possible. Control energy—just as neural activity and its underlying neurophysiology—does not simply deplete, as posited by resource models of control in psychology whose empirical support has eroded [247]. Instead, control energy characterizes continuous readjustments to brain activity by their neural cost, reactively sensitive to inputs such as changing goals or proactively anticipating the demands of the control task in a dynamical prediction feedback loop (“allostasis”) [268, 125, 269, 270, 271, 272].

In place of train traffic causing congestion, interference, and conflict which signal the need for cognitive control, spontaneous activity demands more control energy when it naturally deviates from the control target. As such, cognitive control is not only the consequence of a terminal control target—the activation or deactivation of key brain regions passing some threshold—but is a function of the moment-to-moment processes that guide evolution towards that outcome and continue after that outcome. Network control provides a mathematical model for these continuously evolving functions.

In place of supervision by a metaphorical train conductor or switch operator that follows certain rules and instructions, in network control theory, regions are better network controllers when they mathematically

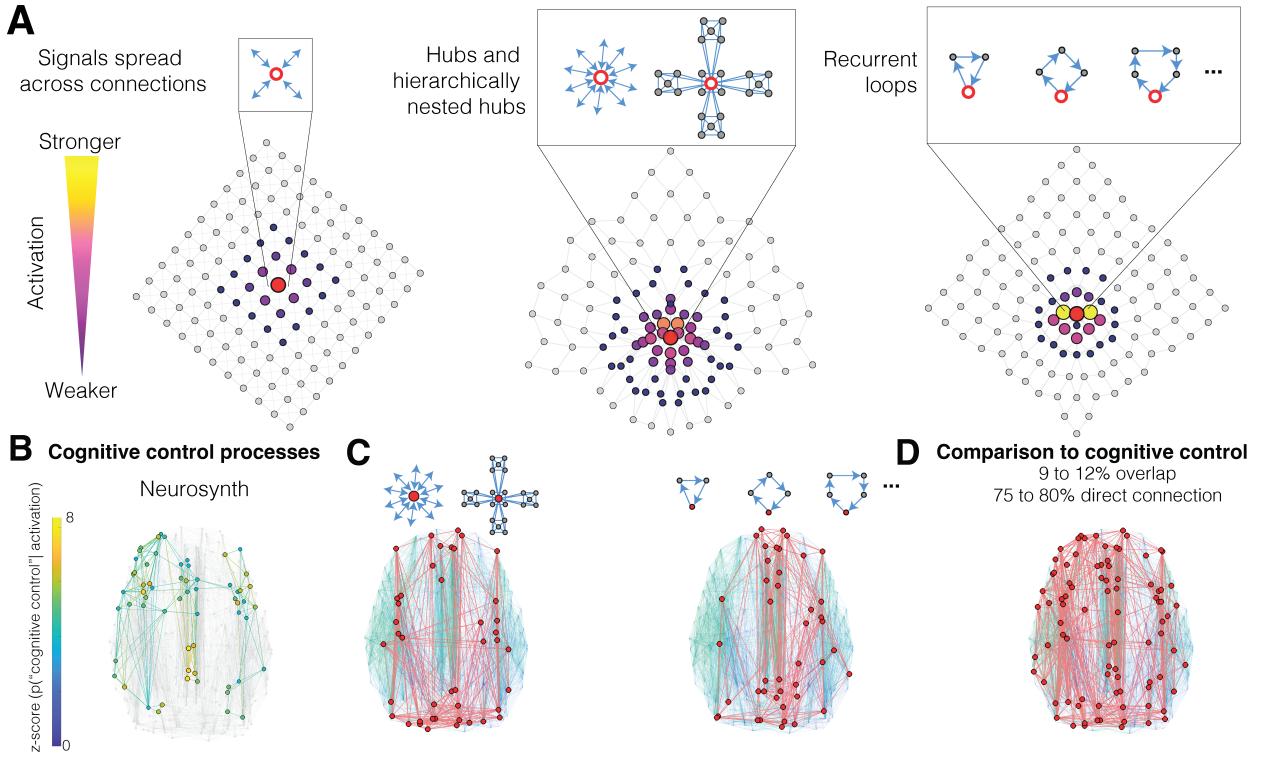


Figure 6: Nested and recurrent organization of networks shape activity flow. **(A)** Linear dynamics makes explicit predictions for how connections influence global signal flow for a given region [261, 262, 263]. For illustration, let us consider different impulse responses, or how different connectivity profiles for a single source region (red node) influences the spread of activation when the red node receives input. Continuous-time dynamics were simulated using the *nctpy* toolbox [95]. *Left:* When the source is embedded with its neighbors in a grid organization, activity spreads evenly. *Middle:* In contrast, if the source region is a hub—having many more connections than its neighbors—then the activity can spread farther, pulling distant nodes inward to be more reachable by using the hub’s many connections. *Right:* If a source region has many more recurrent connections looping back to itself, then activity spreads more than on the grid and in a more concentrated pattern around the source than the hub. Recirculated activity out-competes global spread. These explicit dynamical predictions show how the global activity can be influenced by the differing connectivity profiles of a single source. Network control theory allows us to model multiple simultaneous sources with time-varying input. **(B)** Do regions involved in the cognitive control loops of perception, valuation, and action exhibit recurrent connectivity, suggesting control loops operate at multiple levels? A cognitive control network is visualized, where nodes are highlighted from a Neurosynth meta-analysis of the term “cognitive control” with voxels FDR-corrected at $q < 0.01$ and regions at $z > 3.3$ [273]. Overlaid on the structural connectome, where edges are highlighted if a direct anatomical connection exists between regions of interest. Gray nodes and edges do not pass statistical thresholds. **(C)** The brain has a mixture of connectivity profiles and multiple regions are potential sources of simultaneous activity flow. Highly connected and recurrent hubs most efficiently distribute control inputs to engage control targets [274, 275, 276]. *Left:* Hubs with stronger connections than expected by chance comprise a “rich-club” of about 12% of all regions. *Middle:* The strongest anatomically recurrent hubs comprise about 12% of all regions. **(D)** There is an 89% overlap or direct connection between hubs and the cognitive control nodes, which in combination comprise 27% of brain regions with highly efficient signaling to the other 73% of brain regions. Subcortical regions are not included for illustration but see [274]. See Appendix 4.5 for details.

have more possible, or less energy-intensive, routes to control targets. Given the fact that activity continuously flows between regions at all times, good control regions are points in the network where control energy combines with spontaneous activity to effectively distribute the changes in activity levels towards control targets—a dynamic feedback loop constrained by the physics of the brain.

Convergence and expansion of predicted network control points with prior regions of interest

Intriguingly, while all regions could contribute to this continuous and distributed control process, the theory converges on a similar set of regions as prior investigations due to their advantaged placement for adjusting the flow of activity, including the dorsolateral and ventrolateral prefrontal cortex, superior parietal cortex, dorsal anterior cingulate cortex, thalamus, insula, and frontoparietal network broadly [277, 264, 278, 57, 279, 274, 59] (**Figure 6**). Importantly, the theory also naturally produces graded predictions of other key network control points likely to be involved in the dynamic feedback loop of perception, valuation, and action [274], across brain regions and traditionally segregated psychological functions that have been unconsidered, marginalized, or indirectly linked due to variations of task design. These control points include the orbitofrontal cortex [51], caudate and putamen [280, 281], hippocampus [282, 283, 284, 285], premotor regions [286], default-mode regions like the precuneus or posterior cingulate cortex [228], visual regions [147, 140], and amygdala [212, 287, 288, 120].

Intuitions about cognitive control from the mathematics of network control

Having provided a conceptual overview of network control theory, an important next step is formalizing these core ideas with mathematical concepts. Consider the problem of needing to direct activity in a network to control targets, which are the activity patterns associated with a goal or conflict representation. Each target pattern can be represented as a list of activation values $[t_1, t_2, \dots, t_n]$ for n brain regions. In the context of the railroad analogy, this process is like figuring out how many ways you can arrange ready-made pieces of railroad tracks of fixed length and curvature to reach different target destinations (**Figure 7**). This control problem has multiple solutions constrained by the physics of the brain. Each solution is a different way of factoring one of the target activities t as a summed combination of several repeatable factors x , y , and z , which represent the strength of anatomical connections through which activity flows across. Some target destinations are farther than $x+y+z$, necessitating some repeats of each a , b , and c times, respectively. Taken together, this makes the control problem $t = ax + by + cz$. For example, there are 14 ways of combining the tracks $x = 0.5$, $y = 0.4$, and $z = 0.3$ into the target $t = 3.6$ (**Figure 7A**). The number of possible solutions

to this equation characterizes the probability of control. The quantity $(a + b + c)^2$ characterizes how using a large number of tracks steeply ramps up the energy cost with the squared exponent. Following the previous example, the most efficient solution to factor $t = 3.6$ is to combine x four times with y four times for an energetic cost of $(a + b + c)^2 = (4 + 4 + 0)^2 = 64$. The most expensive solution combines x once, y once, and z nine times for an energetic cost of $(a + b + c)^2 = (1 + 1 + 9)^2 = 121$. Easier and efficient network control has less energetic cost.

Easy and efficient control is less effortful in the sense of marshalling fewer regions and connections to propagate threads of control energy (activity). For easy and efficient control, it is useful to have x , y , and z be large because only a few tracks would be needed to make substantial progress towards the target sum. It is also efficient to have a chain of connections that form strong recurrent pathways because an echo of a previous factor can be recycled in a partially diminished form to advance towards the sum without needing to further increase a , b , or c . The advancement persists, saving progress.

Easy and efficient control further benefits from having additional tracks beyond x , y , and z , such as w . For example, adding just one more train track w into the mix, so that we have four rather than three pieces available, increases the number of solutions from 14 (**Figure 7A**) to 89 (**Figure 7B**). Having additional tracks is characteristic of hierarchically connected hubs, which are highly connected regions that act as nexuses of activity due to widespread connectivity that forms an information backbone for neural communication [289, 290, 291, 275, 68, 262, 292]. Having more tracks allows for a substantially greater repertoire of control solutions and a larger number of efficient possibilities [274, 59], including stabilizing a current state. Such stability emerges not only from efficiency but from being tethered to more neighbors, aggregating activity changes in many directions and muting rapid fluctuations for lower frequency oscillations of activity that may support information integration for external sensorimotor events, as well as internal decision making, motivation and memory processes [293].

In contrast, for difficult and fine-grained control, it is useful to have a diverse range of connection strengths [294]. When x , y , and z are all relatively large, it can create big changes to the sum or a persistent momentum that easily overshoots or strays from the control target pattern of activity t . For a greater number of solutions, it is useful to have some small tracks as well, but this comes at a greater cost to $(a + b + c)^2$ from using small tracks many times. Such small additions are more transient, lacking the hefty structural support of hubs to recycle activity. But what is granted is the ability to flexibly turn to new directions with precision, like making a sharp rather than wide turn when driving a compact car versus a truck. For example, a long track travels farther along before a new change in direction is possible, whereas a short

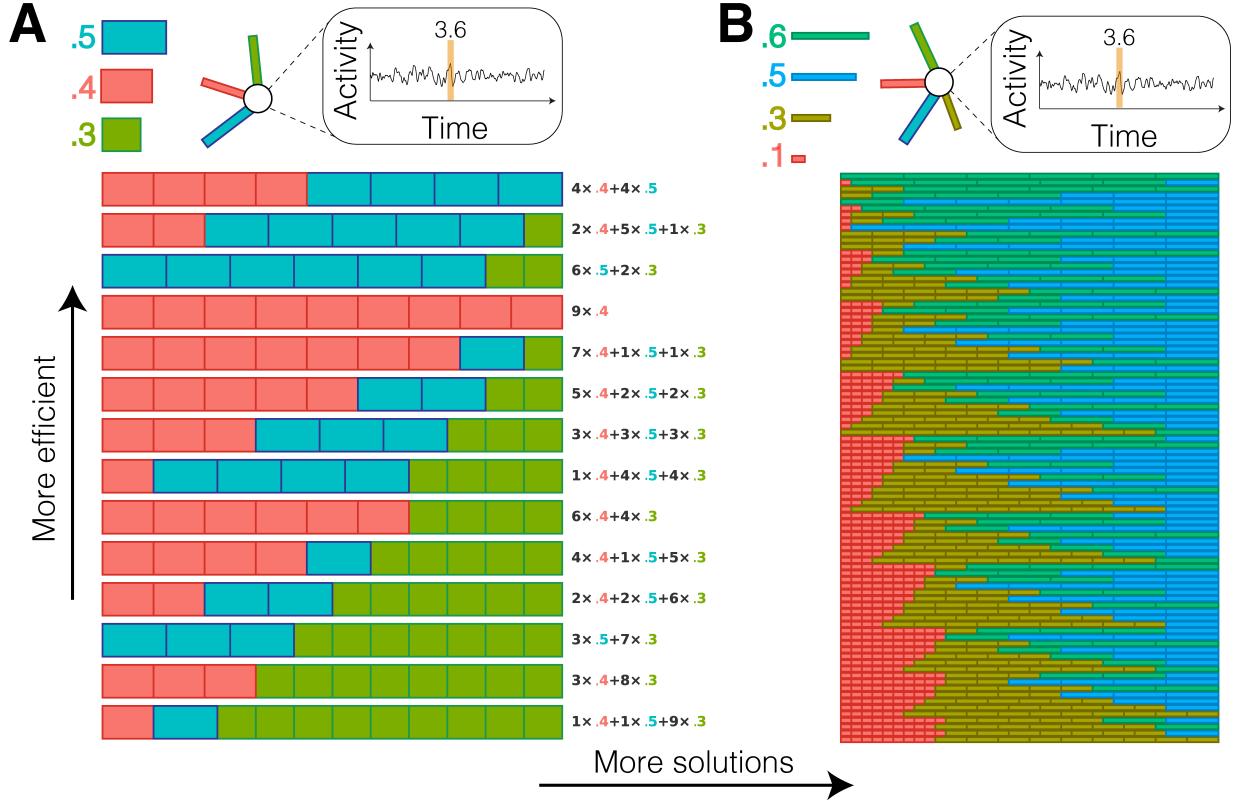


Figure 7: Making train tracks: developing mathematical intuitions about easier and more efficient control. **(A)** Akin to finding all the combinations of railroad tracks in a toy train set that can be used to reach a target destination, the idea underlying the mathematical formalization of control can be illustrated as finding all of the number of ways to make a target pattern (summed activity magnitude of 3.6) with a node (akin to a station) that has three connections of strengths 0.5, 0.4, and 0.3 (where each connection is akin to the specific shape of the railroad track). Each row is 3.6 units of activity wide and specifies one possible solution given the target magnitude and factors. More efficient solutions for reaching the target pattern use fewer connections (akin to the pieces of the track). **(B)** Adding just one more connection vastly increases the repertoire of possible control solutions.

track ends quickly allowing for more frequent changes to target destinations. This responsiveness translates to higher frequency oscillations of activity, contributing to faster and more local processing. Oscillations in different frequency bands may support information integration for memory, learning, and attention [293].

These intuitions show how network control theory defines control based on the number of ways a region can efficiently utilize connections to dynamically guide patterns of activity to functional control targets. Of those possible network control solutions, several useful metrics assume the usage of *at least* the minimal solution theorized to be sufficient for the control problem at hand and others quantify the number of solutions. This illustration is oversimplified in many ways: for example, solutions must simultaneously account for multiple moving targets with multiple stations, the contribution of factors decay at different speeds, some

factors must be used together or in sequence, and solutions are governed by additional spatial and temporal constraints. Nevertheless, the same intuitions apply to more advanced mathematics that allow network control theory to simply but powerfully model how hundreds of regions can simultaneously propagate signals across hundreds of thousands of parallel connectivity pathways over time for any measured or hypothetical control target. Remarkably, as anatomical connections change over development [295, 296] and as control targets reconfigure with learning [274, 297, 264], the brain achieves a unique dynamic reorganization of connectivity and activity for efficient control [294].

There are a number of ways that patterns of activity in a network can be dynamically described by network control theory. This is important, because each pattern has implications for how control is enacted continuously—rather than by a static target pattern of activity—and how it influences the possibility and difficulty of future control such as switching tasks. A control target can be a terminal state that may represent components of a control process, but control theory also determines the persistence and flexibility of the dynamics that lead up to the terminal state and of the dynamics that follow. Persistence is a function of stability granted by low energy, like a ball circling and ultimately resting at the bottom of a bowl. Transience and flexibility is a function of gain granted by high energy, like a ball at the top of a peak ready to roll towards any direction with just the slightest push. The energetic efficiency of dynamics determine the representation’s persistence and flexibility, or how resistant or ready it is to changing inputs. Together, network control characterizes dynamical properties of activity, such as energetic costs, competing activity flow, energy peaks, energy valleys, timescales, and optimality of control routes. These properties open the tantalizing possibility of developing dynamical neural models that closely align with the dynamical properties of behavioral and mental control phenomena, such as different levels of subjective effort, conflict, sensitivity to inputs, stable modes, task switching, flexibility [298, 260, 299, 297, 300, 264].

Applications and evidence for network control theory

Network control theory posits mathematical constructs that have been shown to be associated with psychological and neurobiological control processes, task performance, and neurodevelopment in a number of ways. For example, regions implicated in attention and control processes have been associated with high capacity for network control [260, 279, 277]. Individual differences in network control are correlated with individual differences in performance on cognitive control tasks probing impulsivity, executive function, and working memory [301, 299, 297, 302]. As would be expected if control is costly, top-down activity from frontal regions associated with goal, self-relevance, and control processes uses more control energy than the

bottom-up activity of sensory regions [303]. The amount of cognitive effort demanded by working memory performance and selective attention is related to individual differences in control predicted by network control theory [299, 297, 264], and the network control stability of functional levels of activity for working memory is modulated by dopamine signaling at prefrontal regions [297, 203]. As would be expected if excessively costly control is less likely to happen, transitions to neural states predicted to be difficult due to their costly control energy demands tend to have reduced probability of occurrence [304, 299] and have reduced frequency [305]. As would be expected if control that is costly feels subjectively effortful, neural states predicted to be difficult due to their high control energy involve more demanding cognitive tasks [297, 264]. Finally, the capacity for network control increases with age in conjunction with known periods of neurodevelopment for executive function [295, 299, 277, 296]. Comparatively, humans have a more efficient organization of anatomical connectivity for network control than drosophila and mice [294]

Control energy is a mathematical construct centered on activity change, maintenance, and propagation, and as such, it has been associated with myriad aspects of neural signaling, including dopamine [297], serotonin [306], GABA [307], and metabolic [308, 305] signaling. Network control metrics explain the spatial distribution of the genetic expression of excitatory and inhibitory neurotransmitter receptors that have been proposed to support different forms of cognitive control, including dopaminergic, noradrenergic, cholinergic, and serotonergic receptors, which are required to enact diverse neural dynamics for different cognitive tasks [202, 64, 309, 310, 300]. Together, network control theory offers a mathematical account of cognitive control as an expression of the fundamental architecture of the brain, coordinating the kind, amount, location, and timing of top-down signals to engage activity that is functional for the control tasks at hand. Just as the perception, valuation, and actions of cognitive control operate through dynamic feedback loops, anatomical connectivity itself forms recurrent loops and hierarchically nested hubs that support the efficient propagation of control signals—control loops all the way down.

Limitations and tensions of network control theory

Network control theory has several limitations. First, the theory adopts some limitations of psychological models when incorporating the same control targets as the evoked neural representations of control. These include limitations about what part of a control process the control target is supposed to represent as well as the types of goals embodied in those targets. However, network control metrics are the product of many ways in which activity evolves to a final state, not only of the terminal state itself. The theory is not wedded to any singular terminal target but the dynamics leading to any known and yet unknown targets that can

be multiple realizations of control processes [68, 300].

Second, network control theory is currently more biologically reductive than existing psychological, computational, and neuroscience models [68], but it need not be in the future. Psychological theory speaks more directly to everyday control problems than the neat mathematical solutions for brain dynamics that network control theory presently offers. A preference for functional and algorithmic explanations over ones about neural implementation coheres with the goals of behavioral, cognitive, and computational scientists. Indeed, control behavior is not just the function of deterministic input-output operations of physical machinery [311]. Control behavior also cannot be reduced to a fully specified and task-specific set of coarse goals, finer plans, and granular actions according to iterative feedback [129, 220, 221, 222, 312]. Cognitive control is biopsychosocial, involving the interactions within organisms just as much as with their environment and others [313, 314, 239, 215]. Embracing a plurality of different biopsychosocial goals, which lead to seemingly incommensurable approaches, is where the network approach has great potential [315]. To this end, network control has already been applied across levels of analysis to psychological state and psychopathological symptom networks [316, 317]. Future work could, for example, use multilayer networks to model biological, psychological, and social dynamics generated by existing cognitive computational models within the network control theory approach.

Third, network control theory can seem both too simple by modeling neural dynamics as a linear system and too complex by overlapping with the simpler network metrics of hubs. The assumption of linear dynamics improves interpretability at the cost of biophysical accuracy, allowing for formal predictions about signal flow that can be transparently translated to the additive and multiplicative contributions of particular regions and the strength of their connections in the brain. Linearity makes it mathematically clear, for example, that network control is strongly intertwined with some prior network metrics by definition, such as the communicability of signals based on the strength of anatomical connections and their resulting contribution to the reachability of control targets [318, 319, 274, 59, 275, 276, 320]. Indeed, it has long been argued that control is a function of communication [268, 321, 311], but not that it is merely communication or only a property of communication pathways [322]. Beyond the communicability of control signals, both living organisms and machines regulate behavior towards control targets that change over time through continued interaction and according to diverse contexts. Taken together, control targets and architectures determine the cost of network control, when such costs are excessively large, and precisely what control points contribute to those costs (including hubs but also other key regions often left unconsidered). Network control theory adds conceptual and computational tools to integrate and inform complex spatiotemporal interactions of

measurement, stimulation, and perturbation [98], is predictive of more complex non-linear models of neural dynamics [323, 92, 324], and links the ease of information flow to the efficiency, direction, and localization of control [262, 303]. However, linear assumptions can come at the cost of both biophysical and computational richness [55, 148, 205, 149, 325, 206]. Two promising next steps could be applying network control theory to non-linear computational models in order to provide a complementary explanation for what makes an emerging module of computational units become specialized for control during learning [227, 292, 326] or improving accuracy by using non-linear methods for determining the expected natural flow of activity prior to applying network control theory [327, 328, 329].

3 Opportunities for expanding, formalizing, and integrating cognitive control constructs

How network control theory contributes to the psychological understanding of cognitive control

Network control theory helps synthesize many key ideas in the neuroscience of cognitive control in a precise and mathematically coherent account. For example, it synthesizes the distinction and balance between automatic and controlled activity in service of reaching task-specific activity patterns [55]; the kind, magnitude, and location of guided activation needed for concurrent and parallel control signals to bias top-down and bottom-up systems towards functional target states [56, 57, 58]; the importance of multiple hierarchically structured control hubs [59, 275, 60]; the importance of recurrent connectivity in maintaining stable control representations [123, 126]; the neurodevelopment of control [61, 62, 63]; the physical costs of control [54, 64, 113, 65]; and a general framework for feedback control across levels of analysis [129, 311, 147, 109].

In addition, the extant network control theory literature advances our understanding of cognitive control in several key ways. First, network control theory and conventional brain mapping approaches converge in identifying which regions and networks are most important [279, 277] and how their neurodevelopment may support executive control [279, 295, 301, 296]. Second, several predictions follow from the theory about when and how much control is required given the difficulty of a control task, the probability of activity patterns, resilience to brain damage or disruption, and the effects of perturbations or stimulation [51, 330, 96, 299, 304], providing a useful theoretical testbed for neurostimulation, neurofeedback, and cognitive behavioral strategies. In this way, network-based theories can guide new investigation into regions underappreciated by

prior approaches [59, 330, 96]. Third, the approach is not limited to human psychology, neurobiology, and behavior. Less dependent on the values, brain areas, and functions unobserved in other species, network control can also explain control, cognitive flexibility, attention, and learning in non-human primates [298, 260], mice [331], and *C. elegans* [330, 96]. Fourth, this approach can provide better empirical predictions of behavioral and clinical variability beyond simpler metrics describing local activity and connectivity [330, 332, 333]. For example, network control theory can improve understanding of symptoms across the psychosis spectrum [334, 333, 307, 335], depression [336, 337, 338, 278], attention-deficit/hyperactivity disorder [339], substance use [331, 264], bipolar disorder [340, 341], and epilepsy [259, 308, 342]. Treatments can be modeled and understood using network control theory, including pharmacological interventions [306, 307], neurostimulation [332, 343, 259, 344, 345], and psychological intervention [264, 317]. As such, the approach offers unique conceptual and empirical advantages beyond previous network measures of circuit structure and function [322, 68].

How the psychological understanding of cognitive control contributes to network control theory

The psychology of cognitive control contributes to network control theory new kinds of behaviors that require control, how those behaviors dynamically change, how they are integrated into daily life, and different kinds of control processes or mental representations. In short, psychological approaches reveal new control targets and interpretations of network control's mathematical abstractions.

First, psychology can enrich the interpretation of control theory parameters, such as the time horizon, cost function variables, and the locations of energy input. For example, goals are hierarchical and are construed at differing timescales with differing abstractness versus concreteness [129, 147, 346]. This motivates the design of control inputs to maintain or retrieve goal representations, over long or short time horizons, located at regions that represent more abstract or concrete information. Incongruity between choices and goals can prolong decision processes from conflict and opportunity cost representations [347, 145, 14, 348], suggesting that behavior is guided by cost functions with functional and representational constraints beyond the spatial and energetic ones already built into control theory models [274, 349, 350]. Second, psychological and cognitive neuroscience research has described multiple sources of feedback that can influence the magnitude and location of control inputs, including feedback from regions processing stop signals, expected value, prediction errors, interoceptive signals, arousal, motivational needs, and emotions. This expanded suite of feedback signals could dampen, amplify, or prolong top-down signals from cognitive control and attention

regions [351, 352, 353, 240, 354, 355, 234], and beyond the brain to the body [352, 353, 354, 355]. The regions and neurotransmitters that have been associated with the allocation of control energy overlap with those related to mental fatigue, value, effort, the cost of effort, conflict, and stop signaling [309, 297, 264, 300, 278], suggesting further potential characterizations of control energy [16, 241, 112, 20, 356]. Third, cognitive neuroscience research shows that ongoing spontaneous activity is not merely noise to be overtaken by control input but is often characteristic of default-mode network processes including prospective and episodic memory, thoughts about the self, and social reasoning [357, 358, 359]. New constraints could be added to optimized control trajectories to prolong some functional aspects of prior activity or to minimize the cost of returning. Fourth, goal systems theory, social networks, and new proposals of relational reasoning as a component process of control motivate multilayer network control of interconnected brain regions, goals, needs, motivations, and other people to more directly speak to everyday individual and collective control behaviors [11, 131, 360, 361, 215, 50].

Future directions

“The examples we used came mostly from domains of achievement and instrumental activity. However, the model isn’t just about achievement related affect. Those domains simply provide easy illustrations of the logic. The model is intended to apply to all goal-directed behavior. This includes attempts to attain goals that are amorphous and poorly specified, and for which construing rates of progress isn’t easy. It includes goals for which the idea of rates of progress might at first seem odd. Human goals such as developing and maintaining a sound relationship, being a good [parent], dealing honorably and graciously with acquaintances, seeing someone you care about experience happiness and fulfillment, having a full and rich life, even becoming immersed in the flow of fictional lives in a novel or film are amenable to analysis in these terms. These are all qualities of human experience toward which people try to move, goals that evolve or recur across time, as do most goals underlying human action.”

—Ref. [1]

Toward a new taxonomy of goals: neural bases and conceptual expansion How could any one goal be represented, given a vast universe of possible goals? If goals are attentional templates, task-relevant working memory chunks, or memory schemas, they may not have the rapid flexibility and efficiency necessary to integrate and update a wide variety of changing goals—more like moving targets than checkpoints—which draw from long-term memory, needs, and motivations. Network control theory highlights the theoretical importance for control from multiple regions that represent goals and structured memories, including the

ventromedial prefrontal cortex, orbitofrontal cortex, and hippocampus [362, 363, 364]. These regions may rapidly represent goals as vectors that encode direction, magnitude, timescale, and abstraction [365, 135, 366, 367, 368, 283, 285], in support of both spatial navigation and more abstract goal-oriented behavior. The vectorial targets may be instances of basic needs, motivations, and drives [132, 369, 370, 371] that underlie goal selection [372] and shape signals to regions implementing working memory. Applied to cognitive control, these perspectives echo the “life space”, cognitive map, and vector field metaphor of goals in psychology, with non-spatial elements such as semantic, emotional, social, reward, or motivational features particularly underutilized due to limitations in measurement and modeling tools [365, 135]. Natural language processing and large language models can now be used to represent complex hierarchies of control problems, needs, and goals as a kind of semantic meaning vector [373, 374, 375]. With the tools of network control theory, it is possible to scale beyond the narrow assumption of isolated goals to model the dynamics of more ecologically plausible constellations of interconnected goals and motivations [376, 377, 316, 378, 379].

Social self-regulation and goal-setting through emotion, motivation, mindsets, and narratives. Goals are dynamic rather than static terminal states, and goal pursuit is interwoven with emotion, motivation, mindsets, and narratives, influencing beliefs about what and how goals can be achieved [380, 128, 381, 382, 383, 384, 385, 372, 379, 234, 386, 238]. Beyond an individual’s personal goals, some goals have social origins and goal pursuit can be a collective behavior [361, 215]. Moving beyond static trait measures, dynamic modeling of individual emotional and motivational states using dense real-world ecological momentary assessment can improve our understanding of cognitive control in context [378, 233, 387, 375]. For example, autonomous rather than instructed motivation or goal selection may facilitate goal pursuit and perceived ease by evoking efficient usage of control energy that temporarily stabilizes the focus of the pursuit [379, 372]. This interconnectivity of domains suggests the need for network representations to track the complex influence among different modalities, combining measurements of psychological states from ecological momentary assessment, performance on multiple tasks, and fMRI within a single framework. Multi-layer network models offer a way to simultaneously represent such neural and psychological influences, capturing how goals and emotions co-regulate behavior across different people, timescales, and contexts [361, 360, 316, 317].

Dynamic, biologically constrained computational models of control. Recurrent neural networks and other computational models allow researchers to investigate the learning and adaptation of control strategies across tasks, providing insight into how structural connections adapt and give rise to behavioral outputs [55, 224, 123, 149, 388, 389, 227, 390, 326]. Network control theory could be leveraged to study how the cost of control shapes the geometry of task representations [173, 391], influencing decision-making under

energetic and temporal constraints [98], as well as the conditions that explain why local versus distributed control modules emerge from the communicability of neural networks as they learn [51, 227, 292, 391].

Timeliness of interdigitation. These approaches point toward both theory-rich and ecologically grounded investigations of cognitive control. Now is a momentous period when there is ample neural data at the cellular, genetic, and regional levels, across childhood, adolescence, and in aging, with continuous tracking of behavior, rich methods of disentangling psychological subprocesses using computational models of behavior, and new methods for querying cognitive and emotional states at scale. The observation that this is more data than researchers know what to do with is often used as a call for more theory. Here, we have described one such theory that has accumulated a wide range of evidence in neuroscience, psychology, physics, and bioengineering. We have suggested some possible pathways to bridge the constructs. Towards this end, there are unique opportunities to study cognitive control naturalistically, and to better understand the state dependence thereof in a much more ecologically valid fashion.

Conclusion. The extant network control theory literature undergirds, synthesizes, prunes, and expands cognitive neuroscience theories of control. For example, network control undergirds and synthesizes the role of distributed regions in guided activation in dynamic feedback loops, prunes cultural and moral assumptions about power or the right kinds of goals, and expands the theoretical scope of controllers and control targets relevant to semantic, emotional, social, learning, memory, and motivational processes. Network control theory is within its first decades of investigation, and in the course of expansion and formalization, it may become clear an entirely different model provides a better fit to psychological and neural data. New measurement modalities may revolutionize how we implement state estimation. The limits of connectomic approaches may not bend to theoretical and numerical approaches to modeling. Practical demands for implementation in medical devices or artificial intelligence may necessitate non-linear rather than linear methods. Psychological models of cognitive control may prove more relevant than neural models to the design of personal interventions, scalable measurement, societal goals, structural changes to public policy, economic systems, and cultural norms. Nonetheless, network control theory offers useful tools to intuitively understand cognitive control as a cohesive dynamic and distributed neural process. A major goal of research is to develop multi-level models and biopsychosocial understanding, because these are more robust, capture more data, speak to diverse goals across disciplines, and offer multiple pathways for intervention. In this review, we have shown how psychological and cognitive neuroscience models of control provide the strongest constraints at the behavioral level—by specifying what kinds of behavior count as instances of control and what kinds of goals might one try to pursue—but the weakest constraints on the brain level. By contrast,

network control theory currently provides stronger constraints on the physical level of the brain than it does on the behavioral level. Combining complementary aspects of each could provide the strongest currently available means of theory and modeling across levels of analysis.

Acknowledgments

We thank Dr. Aaron Bornstein for helpful feedback on earlier versions of this manuscript. This research was supported by the Center for Curiosity and was sponsored by the Army Research Office, accomplished under Grant W911NF-18-1-0244. DZ acknowledges support from the George E. Hewitt Foundation for Medical Research. KO was additionally supported by the Center for Brain Mind and Society, ARO grant W911NF-19-1-0411, and grant R01AA023653 from NIAAA. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the US government. The US government is authorized to reproduce and distribute reprints for government purposes notwithstanding any copyright notation herein.

4 Appendix

4.1 Figure 1 Methods

Figure 1 is based on a previously published review in the digital humanities [24], which created a concept network for the psychological construct of self-regulation from linked terms in the American Psychological Association Thesaurus of Psychological Index Terms® [392]. These terms characterize the semantics of psychological science, defining what the words used by psychologists mean to other psychologists [393, 394, 392].

Following the guidance of the procedure detailed in this previously published review [24], we began a search with two terms “cognitive control” and “self-regulation” in the June 2025 edition of the APA Thesaurus [46]. These serve as first-level seed terms. We then obtained the hierarchy of concepts linked to the first-level seed terms to generate a second-level of 13 new seed terms. The second-level terms included: cognitive flexibility, executive function (higher order processes), executive functioning measures, set shifting (attentional set shifting, cognitive set shifting), task switching, agency, emotional regulation (affect regulation, expressive suppression), neurofeedback, self-control (willpower), self-Management, self-monitoring (self-observation), self-monitoring (personality), and self-regulated learning. Finally, we obtained concepts linked to the second-level terms to generate a third-level of 38 terms. This procedure produced 67 terms and their associated definitions (range: 1 to 82 words; mean 19 ± 14 words). The prior work goes one level deeper to fourth-level terms, which is useful because constructs like “goals” and “values” live at that level, but we choose to stop at terms two links away for concision—the sheer breadth of indirectly related constructs dilutes the original meaning of cognitive control or self-regulation. For the few terms without a definition, we used their list of linked terms as a substitute.

Here, our methods diverge from those of the prior paper in order to use recently developed natural language processing and clustering techniques. To cluster the terms and visualize a network of term similarity, we first sought to normalize the number of words used in each term’s definition. To do so, we augment the definitions by a large-language model. Recent work has relied on GPT-4 to write definitions of psychological terms in order to study similarity and ambiguity among constructs [395]. Following their example, we prompt for 100- or 150-word definitions from GPT-4o via the API using the prompt: “You are an expert in psychology and personality theory. Here is a technical dictionary definition of the construct “{label}”: “{definition}”. Expand this into a {length}-word version that retains its scientific accuracy while adding conceptual richness and clarity for a research audience. Use formal psychological language.”

Finally, we applied a topic clustering method using a hierarchical stochastic block model [139]. This method detects community structure in a term-by-definition co-occurrence network [47]. We construct the bipartite network by defining a biadjacency matrix where rows are the psychological terms and columns are the words used in the definitions. Edges represent the number of times a word appears in a term’s definition.

Using a hierarchical stochastic block model, we can cluster definitions and terms into blocks (topics) without needing to predefine the number of topics or rely on a prior probability distribution like in Latent Dirichlet Allocation. Each topic corresponds to a block of words, forming a distribution over words based on their co-occurrence across documents. Similarly, terms are grouped into blocks that share similar word usage, yielding a distribution over terms for each definition word. These distributions emerge naturally from the inferred block structure inferred. Finally, we visualize the term-by-definition bipartite network and use a unique color for each inferred block.

4.2 Figure 2 Methods

Figure 2 uses the data and results from Figures 2, 3, and 4 in Ref. [9]. After downloading their publicly available data, we followed their procedure to process individual performance on 129 task and 64 survey measures ($n = 560$). We used a 560×193 participant-by-measure matrix to create a 193×193 measure-by-measure similarity matrix by calculating the pairwise partial correlation between every pair of measures, controlling for the effect of other measures.

This can be visualized as a fully connected graph, but structure is difficult to ascertain from such a densely connected network. To identify dependencies between the measures, graphical lasso (scikit-learn) applies L1 regularization ($\alpha = 0.15$) to make many entries zero and reduce noisy correlations, making the network more interpretable [396, 397]. This returns a 193×193 measure-by-measure similarity matrix, which we threshold to remove very small correlations $|r| < 0.05$. We obtain an adjacency matrix similar to the one visualized in Figure 2 in Ref. [9]. Each node is a measure’s dependent variable and each edge is the partial correlation between measurement variables.

Here, our methods diverge from those in the prior work [9]. To find and visualize community structure in this sparse graph, we perform clustering by using a stochastic block model implemented in `minimize_nested_blockmodel_dl` by the `graph-tool` library [139]. This community detection approach relies on minimizing a probabilistic description length objective to reveal nested groupings in the network. In contrast, the force-directed layout used in the original figure relies on spatial and visual heuristics, rather than statistical or generative structure. The community detection procedure reveals that the 193×193 matrix is best

modeled by 5 communities, naturally separating survey measurements into two interconnected communities and task measurements into 3 interconnected communities. As in the prior work, these regularization techniques reveal that the task measure communities are largely unrelated to the survey measure communities. However, this may be a result of the application of moderate regularization or the extensive abbreviation of the task durations for the sake of practical experimental implementation, larger sample size, and coverage of measurements.

Finally, we visualize the 193×193 matrix as a network and color its nodes according to how the variable that the node represents loads onto independent latent factors obtained from the exploratory factor analysis reported in Ref. [9]. The 129 task variables loaded onto 5 factors in a 129×5 loading matrix with the following factors: speeded information processing, strategic information processing, perception/response, caution, and discounting. The 64 survey measures loaded onto 12 factors in a 64×12 loading matrix with the following factors: sensation seeking, emotional control, mindfulness, impulsivity, reward sensitivity, goal-directedness, risk perception, eating control, ethical risk-taking, social risk-taking, financial risk-taking, and agreeableness. The task and survey factor matrices contain loadings λ , which characterize the strength and direction of the relationship between the variable and an underlying latent factor.

For each node of the network, we determined the factor with the maximum magnitude loading $|\lambda|$ and color nodes with a unique tone based on their factor names. For each community, we depict the names of the variables corresponding to the community's nodes using word clouds. The size of the words is proportional to its frequency within its community (phrases and names are treated as a single combined word, and then separated after counting their frequency for visualization).

4.3 Figure 3 Methods

To create Figure 3, we obtained bibliometric data by performing a literature search on Web of ScienceTM for the keywords, “cognitive control computational.” This search resulted in 308 articles. We saved the title, keywords, and abstract in BibTeX format.

Using similar methods as for Figure 1, we again performed topic clustering with a hierarchical stochastic block model. The first step is constructing an article-by-word co-occurrence network [47]. This is a biadjacency matrix where rows are articles and columns are the words used in the keywords or abstracts. Edges represent the number of times a word appears in an article.

Several manual and automatic steps were taken for natural language processing. We manually defined and replaced synonymous terms. For example, “addiction” replaced the phrases “substance abuse,” “sub-

stance dependence,” “substance use disorders,” and “craving and relapse.” Known phrases were treated as a single word when calculating frequency. Automatic processing involved stemming and lemmatizing words (e.g., “studies” → “study”), as well as removing stop words (e.g., “the,” “is,” “and,” “in”, and “of”).

Using a hierarchical stochastic block model, we cluster words and articles into blocks (topics). Each topic corresponds to a block of words, forming a distribution over words based on their co-occurrence across articles. Similarly, terms are grouped into blocks that share similar word usage across articles, giving a distribution over terms for each definition word. This data yielded multiple nested layers of communities, where nodes are grouped into blocks at the lowest level, and these blocks themselves are grouped into higher-level blocks. The block hierarchy is depicted as a tree structure in blue. Finally, we visualize the article-by-word bipartite network, using a unique color for each inferred block, and we only visualize the strongest 1000 article-word co-occurrences.

4.4 Figure 5 Methods

To create Figure 5, we reproduced visualizations from Ref. [264] and [262].

4.5 Figure 6 Methods

To create Figure 6, we query Neurosynth for the term “cognitive control” [273]. We download the association test maps, which contain voxels that quantify the probability of cognitive control reported in studies, given activation at specific voxel coordinates. These voxel values are z-scores from a two-way ANOVA testing for the presence of a non-zero association between term use and voxel activation, controlling for multiple comparison with a false discovery rate (FDR) criterion of $q < .01$. We parcellate the voxels into brain regions according to the Glasser atlas [398] and visualize the maximum z-score of the voxels within that region so long as it surpasses a threshold of $z > 3.3$ (corresponding to $p < 0.001$). Hub and recurrence visualizations are reproduced from Refs. [262] and [270].

4.6 Citation Diversity Statement

Recent work in several fields of science has identified a bias in citation practices such that papers from women and other minority scholars are under-cited relative to the number of such papers in the field [399, 400, 401, 402, 403, 404, 405, 406, 407]. Here we sought to proactively consider choosing references that reflect the diversity of the field in thought, form of contribution, gender, race, ethnicity, and other factors. First, we obtained the predicted gender of the first and last author of each reference by using databases that store the

probability of a first name being carried by a woman [403, 408]. By this measure (and excluding self-citations to the first and last authors of our current paper), our references contain 11.07% woman(first)/woman(last), 12.89% man/woman, 16.38% woman/man, and 59.65% man/man. This method is limited in that a) names, pronouns, and social media profiles used to construct the databases may not, in every case, be indicative of gender identity and b) it cannot account for intersex, non-binary, or transgender people. Second, we obtained predicted racial/ethnic category of the first and last author of each reference by databases that store the probability of a first and last name being carried by an author of color [409, 410]. By this measure (and excluding self-citations), our references contain 7.26% author of color (first)/author of color(last), 10.61% white author/author of color, 20.06% author of color/white author, and 62.06% white author/white author. This method is limited in that a) names and Florida Voter Data to make the predictions may not be indicative of racial/ethnic identity, and b) it cannot account for Indigenous and mixed-race authors, or those who may face differential biases due to the ambiguous racialization or ethnicization of their names. We look forward to future work that could help us to better understand how to support equitable practices in science.

References

- [1] C. S. Carver and M. F. Scheier, *On the self-regulation of behavior*. cambridge university press, 2001.
- [2] S. Monsell, “Task switching,” *Trends in cognitive sciences*, vol. 7, no. 3, pp. 134–140, 2003.
- [3] T. Egner and J. Hirsch, “Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information,” *Nature neuroscience*, vol. 8, no. 12, pp. 1784–1790, 2005.
- [4] Y. Munakata, S. A. Herd, C. H. Chatham, B. E. Depue, M. T. Banich, and R. C. O'Reilly, “A unified framework for inhibitory control,” *Trends in cognitive sciences*, vol. 15, no. 10, pp. 453–459, 2011.
- [5] A. Miyake and N. P. Friedman, “The nature and organization of individual differences in executive functions: Four general conclusions,” *Current directions in psychological science*, vol. 21, no. 1, pp. 8–14, 2012.
- [6] S. D. Jamadar, R. Thienel, and F. Karayanidis, “Task switching processes,” *Brain mapping: An encyclopedic reference*, vol. 3, pp. 327–335, 2015.
- [7] L. M. Braunstein, J. J. Gross, and K. N. Ochsner, “Explicit and implicit emotion regulation: a multi-level framework,” *Social cognitive and affective neuroscience*, vol. 12, no. 10, pp. 1545–1557, 2017.
- [8] N. P. Friedman and A. Miyake, “Unity and diversity of executive functions: Individual differences as a window on cognitive structure,” *Cortex*, vol. 86, pp. 186–204, 2017.
- [9] I. W. Eisenberg, P. G. Bissett, A. Zeynep Enkavi, J. Li, D. P. MacKinnon, L. A. Marsch, and R. A. Poldrack, “Uncovering the structure of self-regulation through data-driven ontology discovery,” *Nature communications*, vol. 10, no. 1, p. 2319, 2019.
- [10] R. A. Poldrack, F. W. Sabb, K. Foerde, S. M. Tom, R. F. Asarnow, S. Y. Bookheimer, and B. J. Knowlton, “The neural correlates of motor skill automaticity,” *Journal of Neuroscience*, vol. 25, no. 22, pp. 5356–5364, 2005.
- [11] A. Diamond, “Executive functions,” *Annual review of psychology*, vol. 64, no. 1, pp. 135–168, 2013.
- [12] D. T. Neal, W. Wood, and A. Drolet, “How do people adhere to goals when willpower is low? the profits (and pitfalls) of strong habits.,” *Journal of Personality and Social Psychology*, vol. 104, no. 6, p. 959, 2013.

- [13] T. Amer, K. L. Campbell, and L. Hasher, “Cognitive control as a double-edged sword,” *Trends in cognitive sciences*, vol. 20, no. 12, pp. 905–915, 2016.
- [14] L. Bustamante, F. Lieder, S. Musslick, A. Shenhav, and J. Cohen, “Learning to overexert cognitive control in a stroop task,” *Cognitive, Affective, & Behavioral Neuroscience*, vol. 21, pp. 453–471, 2021.
- [15] D. M. Yee and T. S. Braver, “Interactions of motivation and cognitive control,” *Current opinion in behavioral sciences*, vol. 19, pp. 83–90, 2018.
- [16] A. Shenhav, J. D. Cohen, and M. M. Botvinick, “Dorsal anterior cingulate cortex and the value of control,” *Nature neuroscience*, vol. 19, no. 10, pp. 1286–1291, 2016.
- [17] K. E. Stanovich, “Why humans are cognitive misers and what it means for the great rationality debate,” in *Routledge handbook of bounded rationality*, pp. 196–206, Routledge, 2020.
- [18] A. Touroutoglou, J. Andreano, B. C. Dickerson, and L. F. Barrett, “The tenacious brain: How the anterior mid-cingulate contributes to achieving goals,” *Cortex*, vol. 123, pp. 12–29, 2020.
- [19] T. Ribot, “Voluntary attention.,” 1911.
- [20] W. Kool and M. Botvinick, “Mental labour,” *Nature human behaviour*, vol. 2, no. 12, pp. 899–908, 2018.
- [21] A. L. Duckworth and M. L. Kern, “A meta-analysis of the convergent validity of self-control measures,” *Journal of research in personality*, vol. 45, no. 3, pp. 259–268, 2011.
- [22] W. Mischel, O. Ayduk, M. G. Berman, B. Casey, I. H. Gotlib, J. Jonides, E. Kross, T. Teslovich, N. L. Wilson, V. Zayas, *et al.*, “willpower’over the life span: decomposing self-regulation,” *Social cognitive and affective neuroscience*, vol. 6, no. 2, pp. 252–256, 2011.
- [23] K. Fujita, “On conceptualizing self-control as more than the effortful inhibition of impulses,” *Personality and social psychology review*, vol. 15, no. 4, pp. 352–366, 2011.
- [24] J. T. Burman, C. D. Green, and S. Shanker, “On the meanings of self-regulation: Digital humanities in service of conceptual clarity,” *Child development*, vol. 86, no. 5, pp. 1507–1521, 2015.
- [25] J. T. Nigg, “Annual research review: On the relations among self-regulation, self-control, executive functioning, effortful control, cognitive control, impulsivity, risk-taking, and inhibition for developmental psychopathology,” *Journal of child psychology and psychiatry*, vol. 58, no. 4, pp. 361–383, 2017.

- [26] M. Inzlicht, K. M. Werner, J. L. Briskin, and B. W. Roberts, “Integrating models of self-regulation,” *Annual review of psychology*, vol. 72, no. 1, pp. 319–345, 2021.
- [27] A. F. Arnsten and K. Rubia, “Neurobiological circuits regulating attention, cognitive control, motivation, and emotion: disruptions in neurodevelopmental psychiatric disorders,” *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 51, no. 4, pp. 356–367, 2012.
- [28] L. M. McTeague, J. Huemer, D. M. Carreon, Y. Jiang, S. B. Eickhoff, and A. Etkin, “Identification of common neural circuit disruptions in cognitive control across psychiatric disorders,” *American Journal of Psychiatry*, vol. 174, no. 7, pp. 676–685, 2017.
- [29] D. M. Barch, A. Culbreth, and J. Sheffield, “Systems level modeling of cognitive control in psychiatric disorders: A focus on schizophrenia,” in *Computational psychiatry*, pp. 145–173, Elsevier, 2018.
- [30] A. L. Romer, A. R. Hariri, and T. J. Strauman, “Regulatory focus and the p factor: Evidence for self-regulatory dysfunction as a transdiagnostic feature of general psychopathology,” *Journal of Psychiatric Research*, vol. 137, pp. 178–185, 2021.
- [31] N. Park, C. Peterson, and M. E. Seligman, “Strengths of character and well-being,” *Journal of social and Clinical Psychology*, vol. 23, no. 5, pp. 603–619, 2004.
- [32] A. L. Duckworth, C. Peterson, M. D. Matthews, and D. R. Kelly, “Grit: perseverance and passion for long-term goals.,” *Journal of personality and social psychology*, vol. 92, no. 6, p. 1087, 2007.
- [33] W. Hofmann, P. Meindl, M. Mooijman, and J. Graham, “Morality and self-control: How they are intertwined and where they differ,” *Current Directions in Psychological Science*, vol. 27, no. 4, pp. 286–291, 2018.
- [34] A. Wente, X. Zhao, A. Gopnik, C. Kang, and T. Kushnir, “The developmental and cultural origins of our beliefs about self-control,” *Surrounding self-control*, 2020.
- [35] S. P. Speer, A. Smidts, and M. A. Boksem, “Cognitive control and dishonesty,” *Trends in Cognitive Sciences*, vol. 26, no. 9, pp. 796–808, 2022.
- [36] U. Müller and K. Kerns, “The development of executive function,” *Handbook of child psychology and developmental science: Cognitive processes*, vol. 2, pp. 571–623, 2015.

- [37] B. Luna, “The relevance of immaturities in the juvenile brain to culpability and rehabilitation,” *The Hastings law journal*, vol. 63, no. 6, p. 1469, 2012.
- [38] B. Luna and C. Wright, “Adolescent brain development: Implications for the juvenile criminal justice system.,” 2016.
- [39] A. L. Duckworth and M. E. Seligman, “Self-discipline outdoes iq in predicting academic performance of adolescents,” *Psychological science*, vol. 16, no. 12, pp. 939–944, 2005.
- [40] C. Blair and R. P. Razza, “Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten,” *Child development*, vol. 78, no. 2, pp. 647–663, 2007.
- [41] T. F. Heatherton and D. D. Wagner, “Cognitive neuroscience of self-regulation failure,” *Trends in cognitive sciences*, vol. 15, no. 3, pp. 132–139, 2011.
- [42] T. E. Moffitt, L. Arseneault, D. Belsky, N. Dickson, R. J. Hancox, H. Harrington, R. Houts, R. Poulton, B. W. Roberts, S. Ross, *et al.*, “A gradient of childhood self-control predicts health, wealth, and public safety,” *Proceedings of the national Academy of Sciences*, vol. 108, no. 7, pp. 2693–2698, 2011.
- [43] A. Gennara, J. Peetz, and M. Milyavskaya, “When more is less: Self-control strategies are seen as less indicative of self-control than just willpower,” *Journal of Experimental Social Psychology*, vol. 106, p. 104457, 2023.
- [44] D. M. Lydon, S. J. Wilson, A. Child, and C. F. Geier, “Adolescent brain maturation and smoking: what we know and where we’re headed,” *Neuroscience & Biobehavioral Reviews*, vol. 45, pp. 323–342, 2014.
- [45] J. P. Bermúdez and S. Murray, “Believe in your self-control: Lay theories of self-control and their downstream effects,” *Current Opinion in Psychology*, vol. 60, p. 101879, 2024.
- [46] American Psychological Association, “Apa thesaurus of psychological index terms.” <https://www.apa.org/pubs/databases/training/thesaurus>, 2023. Accessed via PsycINFO; provides controlled vocabulary for psychological concepts.
- [47] M. Gerlach, T. P. Peixoto, and E. G. Altmann, “A network approach to topic models,” *Science advances*, vol. 4, no. 7, p. eaaq1360, 2018.

- [48] H. P. Kotabe and W. Hofmann, “On integrating the components of self-control,” *Perspectives on Psychological Science*, vol. 10, no. 5, pp. 618–638, 2015.
- [49] M. Milyavskaya, E. T. Berkman, and D. T. De Ridder, “The many faces of self-control: Tacit assumptions and recommendations to deal with them,” *Motivation Science*, vol. 5, no. 1, p. 79, 2019.
- [50] S. A. Bunge, “How should we slice up the executive function pie? striving toward an ontology of cognitive control processes,” *Mind, Brain, and Education*, vol. 18, no. 1, pp. 17–27, 2024.
- [51] B. R. Eisenreich, R. Akaishi, and B. Y. Hayden, “Control without controllers: Toward a distributed neuroscience of executive control,” *Journal of cognitive neuroscience*, vol. 29, no. 10, pp. 1684–1698, 2017.
- [52] L. Pessoa, “The entangled brain,” *Journal of Cognitive Neuroscience*, vol. 35, no. 3, pp. 349–360, 2023.
- [53] S. Noble, J. Curtiss, L. Pessoa, and D. Scheinost, “The tip of the iceberg: a call to embrace anti-localizationism in human neuroscience research,” 2023.
- [54] A. Shenhav, M. M. Botvinick, and J. D. Cohen, “The expected value of control: an integrative theory of anterior cingulate cortex function,” *Neuron*, vol. 79, no. 2, pp. 217–240, 2013.
- [55] J. D. Cohen, K. Dunbar, and J. L. McClelland, “On the control of automatic processes: a parallel distributed processing account of the stroop effect.,” *Psychological review*, vol. 97, no. 3, p. 332, 1990.
- [56] R. Desimone, J. Duncan, *et al.*, “Neural mechanisms of selective visual attention,” *Annual review of neuroscience*, vol. 18, no. 1, pp. 193–222, 1995.
- [57] E. K. Miller, “The prefrontal cortex and cognitive control,” *Nature reviews neuroscience*, vol. 1, no. 1, pp. 59–65, 2000.
- [58] D. Badre, “Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes,” *Trends in cognitive sciences*, vol. 12, no. 5, pp. 193–200, 2008.
- [59] M. W. Cole, “Control and connectivity: Dynamic networks in the human brain,” *The Wiley Handbook of Cognitive Control*, pp. 314–333, 2017.
- [60] D. Badre and D. E. Nee, “Frontal cortex and the hierarchical control of behavior,” *Trends in cognitive sciences*, vol. 22, no. 2, pp. 170–188, 2018.

- [61] B. Luna, S. Marek, B. Larsen, B. Tervo-Clemmens, and R. Chahal, “An integrative model of the maturation of cognitive control,” *Annual review of neuroscience*, vol. 38, pp. 151–170, 2015.
- [62] S. Marek, K. Hwang, W. Foran, M. N. Hallquist, and B. Luna, “The contribution of network organization and integration to the development of cognitive control,” *PLoS biology*, vol. 13, no. 12, p. e1002328, 2015.
- [63] E. A. Crone and N. Steinbeis, “Neural perspectives on cognitive control development during childhood and adolescence,” *Trends in cognitive sciences*, vol. 21, no. 3, pp. 205–215, 2017.
- [64] R. Cools, “The cost of dopamine for dynamic cognitive control,” *Current Opinion in Behavioral Sciences*, vol. 4, pp. 152–159, 2015.
- [65] A. Westbrook, R. Van Den Bosch, J. Määttä, L. Hofmans, D. Papadopetraki, R. Cools, and M. Frank, “Dopamine promotes cognitive effort by biasing the benefits versus costs of cognitive work,” *Science*, vol. 367, no. 6484, pp. 1362–1366, 2020.
- [66] J. D. Medaglia, “Clarifying cognitive control and the controllable connectome,” *Wiley Interdisciplinary Reviews: Cognitive Science*, vol. 10, no. 1, p. e1471, 2019.
- [67] C. W. Lynn and D. S. Bassett, “The physics of brain network structure, function and control,” *Nature Reviews Physics*, vol. 1, no. 5, pp. 318–332, 2019.
- [68] D. Badre, “Brain networks for cognitive control: four unresolved questions,” *Intrusive Thinking across Neuropsychiatric Disorders: From Molecules to Free Will*, pp. 203–228, 2020.
- [69] P. M. Gollwitzer, “The history of the concept of goals,” *A Century of Psychology (Psychology Revivals)*, pp. 224–240, 2013.
- [70] V. Brandstätter and K. Bernecker, “Persistence and disengagement in personal goal pursuit,” *Annual review of psychology*, vol. 73, no. 1, pp. 271–299, 2022.
- [71] D. M. Yee and T. S. Braver, “20 neurocomputational models of cognitive control,”
- [72] D. Yee and T. Braver, “Computational models of cognitive control: past and current approaches,” 2018.
- [73] A. B. Satpute, K. N. Ochsner, and D. Badre, “The neuroscience of goal-directed behavior,” in *Goal-directed behavior*, pp. 49–84, Psychology Press, 2012.

- [74] T. Egner, *The Wiley handbook of cognitive control*. John Wiley & Sons, 2017.
- [75] G. Gratton, P. Cooper, M. Fabiani, C. S. Carter, and F. Karayanidis, “Dynamics of cognitive control: Theoretical bases, paradigms, and a view for the future,” *Psychophysiology*, vol. 55, no. 3, p. e13016, 2018.
- [76] D. Badre, *On task: how our brain gets things done*. Princeton University Press, 2022.
- [77] N. P. Friedman and T. W. Robbins, “The role of prefrontal cortex in cognitive control and executive function,” *Neuropsychopharmacology*, vol. 47, no. 1, pp. 72–89, 2022.
- [78] V. Menon and M. D’Esposito, “The role of pfc networks in cognitive control and executive function,” *Neuropsychopharmacology*, vol. 47, no. 1, pp. 90–103, 2022.
- [79] N. Clairis and A. Lopez-Persem, “Debates on the dorsomedial prefrontal/dorsal anterior cingulate cortex: insights for future research,” *Brain*, vol. 146, no. 12, pp. 4826–4844, 2023.
- [80] V. Menon and L. Q. Uddin, “Saliency, switching, attention and control: a network model of insula function,” *Brain structure and function*, vol. 214, pp. 655–667, 2010.
- [81] L. Q. Uddin, “Cognitive and behavioural flexibility: neural mechanisms and clinical considerations,” *Nature Reviews Neuroscience*, vol. 22, no. 3, pp. 167–179, 2021.
- [82] N. Silvestrini, S. Musslick, A. S. Berry, and E. Vassena, “An integrative effort: Bridging motivational intensity theory and recent neurocomputational and neuronal models of effort and control allocation.,” *Psychological Review*, vol. 130, no. 4, p. 1081, 2023.
- [83] T. A. Niendam, A. R. Laird, K. L. Ray, Y. M. Dean, D. C. Glahn, and C. S. Carter, “Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions,” *Cognitive, Affective, & Behavioral Neuroscience*, vol. 12, pp. 241–268, 2012.
- [84] L. H. Somerville and B. Casey, “Developmental neurobiology of cognitive control and motivational systems,” *Current opinion in neurobiology*, vol. 20, no. 2, pp. 236–241, 2010.
- [85] B. Luna, A. Padmanabhan, and K. O’Hearn, “What has fmri told us about the development of cognitive control through adolescence?,” *Brain and cognition*, vol. 72, no. 1, pp. 101–113, 2010.
- [86] Y. Munakata, H. R. Snyder, and C. H. Chatham, “Developing cognitive control: Three key transitions,” *Current directions in psychological science*, vol. 21, no. 2, pp. 71–77, 2012.

- [87] A. E. Motter, “Networkcontrology,” *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 25, no. 9, 2015.
- [88] Y.-Y. Liu and A.-L. Barabási, “Control principles of complex systems,” *Reviews of Modern Physics*, vol. 88, no. 3, p. 035006, 2016.
- [89] R. M. D’Souza, M. Di Bernardo, and Y.-Y. Liu, “Controlling complex networks with complex nodes,” *Nature Reviews Physics*, vol. 5, no. 4, pp. 250–262, 2023.
- [90] E. Tang and D. S. Bassett, “Colloquium: Control of dynamics in brain networks,” *Reviews of modern physics*, vol. 90, no. 3, p. 031003, 2018.
- [91] P. Srivastava, E. Nozari, J. Z. Kim, H. Ju, D. Zhou, C. Becker, F. Pasqualetti, G. J. Pappas, and D. S. Bassett, “Models of communication and control for brain networks: distinctions, convergence, and future outlook,” *Network Neuroscience*, vol. 4, no. 4, pp. 1122–1159, 2020.
- [92] J. Z. Kim and D. S. Bassett, “Linear dynamics and control of brain networks,” *Neural Engineering*, pp. 497–518, 2020.
- [93] S. L. Brunton and J. N. Kutz, *Data-driven science and engineering: Machine learning, dynamical systems, and control*. Cambridge University Press, 2022.
- [94] T. M. Karrer, J. Z. Kim, J. Stiso, A. E. Kahn, F. Pasqualetti, U. Habel, and D. S. Bassett, “A practical guide to methodological considerations in the controllability of structural brain networks,” *Journal of neural engineering*, vol. 17, no. 2, p. 026031, 2020.
- [95] L. Parkes, J. Z. Kim, J. Stiso, J. K. Brynildsen, M. Cieslak, S. Covitz, R. E. Gur, R. C. Gur, F. Pasqualetti, R. T. Shinohara, *et al.*, “Using network control theory to study the dynamics of the structural connectome,” *bioRxiv*, pp. 2023–08, 2023.
- [96] E. K. Towlson, P. E. Vértes, G. Yan, Y. L. Chew, D. S. Walker, W. R. Schafer, and A.-L. Barabási, “*Caenorhabditis elegans* and the network control framework—faqs,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 373, no. 1758, p. 20170372, 2018.
- [97] J. D. Medaglia, P. Zurn, W. Sinnott-Armstrong, and D. S. Bassett, “Mind control as a guide for the mind,” *Nature Human Behaviour*, vol. 1, no. 6, p. 0119, 2017.

- [98] K. V. Shenoy and J. C. Kao, “Measurement, manipulation and modeling of brain-wide neural population dynamics,” *Nature communications*, vol. 12, no. 1, p. 633, 2021.
- [99] P. Srivastava, P. Fotiadis, L. Parkes, and D. S. Bassett, “The expanding horizons of network neuroscience: From description to prediction and control,” *Neuroimage*, vol. 258, p. 119250, 2022.
- [100] W. James, “The principles of,” *Psychology*, vol. 2, p. 94, 1890.
- [101] D. Davidson, “How is weakness of the will possible?,” *Essays on actions and events*, pp. 21–42, 2001.
- [102] P. Plato, *Phaedrus*, vol. 1. Bobbs-Merrill Indianapolis, 1952.
- [103] N. Elias, “The civilizing process,” 1939.
- [104] D. Read, G. Loewenstein, and S. Kalyanaraman, “Mixing virtue and vice: Combining the immediacy effect and the diversification heuristic,” *Journal of Behavioral Decision Making*, vol. 12, no. 4, pp. 257–273, 1999.
- [105] S. Pinker, *The better angels of our nature: Why violence has declined*. Penguin books, 2012.
- [106] H. Heckhausen, “Historical trends in motivation research,” in *Motivation and action*, pp. 15–65, Springer, 2018.
- [107] D. E. Nee, T. D. Wager, and J. Jonides, “Interference resolution: insights from a meta-analysis of neuroimaging tasks,” *Cognitive, Affective, & Behavioral Neuroscience*, vol. 7, no. 1, pp. 1–17, 2007.
- [108] D. E. Nee and J. Jonides, “Dissociable interference-control processes in perception and memory,” *Psychological science*, vol. 19, no. 5, pp. 490–500, 2008.
- [109] G. A. Miller, E. Galanter, and K. H. Pribram, *Plans and the Structure of Behavior*. Henry Holt and Co., 1960.
- [110] N. D. Daw, Y. Niv, and P. Dayan, “Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control,” *Nature neuroscience*, vol. 8, no. 12, pp. 1704–1711, 2005.
- [111] J. T. McGuire and J. W. Kable, “Rational temporal predictions can underlie apparent failures to delay gratification.,” *Psychological review*, vol. 120, no. 2, p. 395, 2013.
- [112] E. T. Berkman, C. A. Hutcherson, J. L. Livingston, L. E. Kahn, and M. Inzlicht, “Self-control as value-based choice,” *Current directions in psychological science*, vol. 26, no. 5, pp. 422–428, 2017.

- [113] F. Lieder, A. Shenhav, S. Musslick, and T. L. Griffiths, “Rational metareasoning and the plasticity of cognitive control,” *PLoS computational biology*, vol. 14, no. 4, p. e1006043, 2018.
- [114] S. Musslick and J. D. Cohen, “Rationalizing constraints on the capacity for cognitive control,” *Trends in cognitive sciences*, vol. 25, no. 9, pp. 757–775, 2021.
- [115] S. Padmala and L. Pessoa, “Reward reduces conflict by enhancing attentional control and biasing visual cortical processing,” *Journal of cognitive neuroscience*, vol. 23, no. 11, pp. 3419–3432, 2011.
- [116] W. Kool, A. Shenhav, and M. M. Botvinick, “Cognitive control as cost-benefit decision making,” *The Wiley handbook of cognitive control*, pp. 167–189, 2017.
- [117] A. J. Shackman, T. V. Salomons, H. A. Slagter, A. S. Fox, J. J. Winter, and R. J. Davidson, “The integration of negative affect, pain and cognitive control in the cingulate cortex,” *Nature Reviews Neuroscience*, vol. 12, no. 3, pp. 154–167, 2011.
- [118] P. D. Zelazo and S. M. Carlson, “Hot and cool executive function in childhood and adolescence: Development and plasticity,” *Child development perspectives*, vol. 6, no. 4, pp. 354–360, 2012.
- [119] K. N. Ochsner, J. A. Silvers, and J. T. Buhle, “Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion,” *Annals of the New York Academy of Sciences*, vol. 1251, no. 1, pp. E1–E24, 2012.
- [120] L. Pessoa, “Cognitive control and emotional processing,” *The Wiley handbook of cognitive control*, pp. 392–407, 2017.
- [121] M. Enisman and T. Kleiman, “The relative difficulty of resolving motivational conflicts is affective context-dependent.,” *Emotion*, 2024.
- [122] W. Wood, A. Mazar, and D. T. Neal, “Habits and goals in human behavior: Separate but interacting systems,” *Perspectives on Psychological Science*, vol. 17, no. 2, pp. 590–605, 2022.
- [123] N. P. Rougier, D. C. Noelle, T. S. Braver, J. D. Cohen, and R. C. O'Reilly, “Prefrontal cortex and flexible cognitive control: Rules without symbols,” *Proceedings of the National Academy of Sciences*, vol. 102, no. 20, pp. 7338–7343, 2005.
- [124] E. Koechlin and C. Summerfield, “An information theoretical approach to prefrontal executive function,” *Trends in cognitive sciences*, vol. 11, no. 6, pp. 229–235, 2007.

- [125] T. S. Braver, “The variable nature of cognitive control: a dual mechanisms framework,” *Trends in cognitive sciences*, vol. 16, no. 2, pp. 106–113, 2012.
- [126] M. M. Botvinick and J. D. Cohen, “The computational and neural basis of cognitive control: Charted territory and new frontiers,” *Cognitive science*, vol. 38, no. 6, pp. 1249–1285, 2014.
- [127] M. Botvinick and T. Braver, “Motivation and cognitive control: from behavior to neural mechanism,” *Annual review of psychology*, vol. 66, no. 1, pp. 83–113, 2015.
- [128] E. L. Deci and R. M. Ryan, “The” what” and” why” of goal pursuits: Human needs and the self-determination of behavior,” *Psychological inquiry*, vol. 11, no. 4, pp. 227–268, 2000.
- [129] C. S. Carver and M. F. Scheier, “Control theory: A useful conceptual framework for personality–social, clinical, and health psychology.,” *Psychological bulletin*, vol. 92, no. 1, p. 111, 1982.
- [130] A. W. Kruglanski, J. J. Bélanger, X. Chen, C. Köpetz, A. Pierro, and L. Mannetti, “The energetics of motivated cognition: a force-field analysis.,” *Psychological review*, vol. 119, no. 1, p. 1, 2012.
- [131] A. W. Kruglanski, M. Chernikova, M. Babush, M. Dugas, and B. M. Schumpe, “The architecture of goal systems: Multifinality, equifinality, and counterfinality in means—end relations,” in *Advances in motivation science*, vol. 2, pp. 69–98, Elsevier, 2015.
- [132] J. W. Atkinson, “Motivational determinants of risk-taking behavior.,” *Psychological review*, vol. 64, no. 6p1, p. 359, 1957.
- [133] J. Shah and E. T. Higgins, “Expectancy× value effects: Regulatory focus as determinant of magnitude and direction.,” *Journal of personality and social psychology*, vol. 73, no. 3, p. 447, 1997.
- [134] N. T. Feather, *Expectations and actions: Expectancy-value models in psychology*. Routledge, 1982.
- [135] K. Lewin, *Principles of topological psychology*. Read Books Ltd, 1936.
- [136] R. M. Ryan, E. L. Deci, *et al.*, “Overview of self-determination theory: An organismic dialectical perspective,” *Handbook of self-determination research*, vol. 2, no. 3-33, p. 36, 2002.
- [137] P. W. Glimcher and E. Fehr, *Neuroeconomics: Decision making and the brain*. Academic Press, 2013.
- [138] Y.-L. Boureau, P. Sokol-Hessner, and N. D. Daw, “Deciding how to decide: Self-control and meta-decision making,” *Trends in cognitive sciences*, vol. 19, no. 11, pp. 700–710, 2015.

- [139] T. P. Peixoto, “The graph-tool python library,” *figshare*, 2014.
- [140] J. M. Fuster, “Prefrontal cortex and the bridging of temporal gaps in the perception-action cycle.,” *Annals of the New York Academy of Sciences*, vol. 608, pp. 318–29, 1990.
- [141] K. N. Ochsner and J. J. Gross, “The neural bases of emotion and emotion regulation: A valuation perspective.,” 2014.
- [142] J. J. Gross, “Emotion regulation: Current status and future prospects,” *Psychological inquiry*, vol. 26, no. 1, pp. 1–26, 2015.
- [143] M. Rmus, S. D. McDougle, and A. G. Collins, “The role of executive function in shaping reinforcement learning,” *Current Opinion in Behavioral Sciences*, vol. 38, pp. 66–73, 2021.
- [144] G. A. Miller, G. Eugene, and K. H. Pribram, *Plans and the Structure of Behaviour*. Routledge, 2017.
- [145] R. Frömer, C. K. Dean Wolf, and A. Shenhav, “Goal congruency dominates reward value in accounting for behavioral and neural correlates of value-based decision-making,” *Nature communications*, vol. 10, no. 1, p. 4926, 2019.
- [146] S. D. McDougle, I. C. Ballard, B. Baribault, S. J. Bishop, and A. G. Collins, “Executive function assigns value to novel goal-congruent outcomes,” *Cerebral Cortex*, vol. 32, no. 1, pp. 231–247, 2022.
- [147] W. T. Powers and W. T. Powers, “Behavior: The control of perception,” 1973.
- [148] M. M. Botvinick, T. S. Braver, D. M. Barch, C. S. Carter, and J. D. Cohen, “Conflict monitoring and cognitive control.,” *Psychological review*, vol. 108, no. 3, p. 624, 2001.
- [149] S. A. Herd, T. E. Hazy, C. H. Chatham, A. M. Brant, N. P. Friedman, *et al.*, “A neural network model of individual differences in task switching abilities,” *Neuropsychologia*, vol. 62, pp. 375–389, 2014.
- [150] H. Ritz, X. Leng, and A. Shenhav, “Cognitive control as a multivariate optimization problem,” *Journal of Cognitive Neuroscience*, vol. 34, no. 4, pp. 569–591, 2022.
- [151] D. C. Park and P. Reuter-Lorenz, “The adaptive brain: aging and neurocognitive scaffolding,” *Annual review of psychology*, vol. 60, pp. 173–196, 2009.
- [152] J. Parvizi, V. Rangarajan, W. R. Shirer, N. Desai, and M. D. Greicius, “The will to persevere induced by electrical stimulation of the human cingulate gyrus,” *Neuron*, vol. 80, no. 6, pp. 1359–1367, 2013.

- [153] P. A. Reuter-Lorenz and D. C. Park, “How does it stack up? revisiting the scaffolding theory of aging and cognition,” *Neuropsychology review*, vol. 24, no. 3, pp. 355–370, 2014.
- [154] J. F. Cavanagh and M. J. Frank, “Frontal theta as a mechanism for cognitive control,” *Trends in cognitive sciences*, vol. 18, no. 8, pp. 414–421, 2014.
- [155] A. Widge, S. Zorowitz, I. Basu, A. Paulk, S. Cash, E. Eskandar, T. Deckersbach, E. Miller, and D. Dougherty, “Deep brain stimulation of the internal capsule enhances human cognitive control and prefrontal cortex function,” *Nature communications*, vol. 10, no. 1, p. 1536, 2019.
- [156] A. Livnat and N. Pippenger, “An optimal brain can be composed of conflicting agents,” *Proceedings of the National Academy of Sciences*, vol. 103, no. 9, pp. 3198–3202, 2006.
- [157] H. Ritz and A. Shenhav, “Orthogonal neural encoding of targets and distractors supports multivariate cognitive control,” *Nature human behaviour*, vol. 8, no. 5, pp. 945–961, 2024.
- [158] L. K. Held, L. Vermeylen, D. Dignath, W. Notebaert, R. M. Krebs, and S. Braem, “Reinforcement learning of adaptive control strategies,” *Communications Psychology*, vol. 2, no. 1, p. 8, 2024.
- [159] J. Fritz and G. Dreisbach, “Conflicts as aversive signals: Conflict priming increases negative judgments for neutral stimuli,” *Cognitive, Affective, & Behavioral Neuroscience*, vol. 13, no. 2, pp. 311–317, 2013.
- [160] N. S. Narayanan, J. F. Cavanagh, M. J. Frank, and M. Laubach, “Common medial frontal mechanisms of adaptive control in humans and rodents,” *Nature neuroscience*, vol. 16, no. 12, pp. 1888–1895, 2013.
- [161] J. M. Fuster and G. E. Alexander, “Neuron activity related to short-term memory,” *Science*, vol. 173, no. 3997, pp. 652–654, 1971.
- [162] A. Diamond and P. Goldman-Rakic, “Comparative development in human infants and infant rhesus monkeys of cognitive functions that depend on prefrontal cortex,” in *Society for Neuroscience Abstracts*, vol. 12, p. 2, 1986.
- [163] P. Goldman-Rakic, “S.(1987). circuitry of primate prefrontal cortex and regulation of behaviour by representational memory,” *Handbook of physiology: the nervous system (ed. F. Plum)*, pp. 373–417, 1987.
- [164] S. Funahashi, M. V. Chafee, and P. S. Goldman-Rakic, “Prefrontal neuronal activity in rhesus monkeys performing a delayed anti-saccade task,” *Nature*, vol. 365, no. 6448, pp. 753–756, 1993.

- [165] W. F. Asaad, G. Rainer, and E. K. Miller, “Task-specific neural activity in the primate prefrontal cortex,” *Journal of neurophysiology*, vol. 84, no. 1, pp. 451–459, 2000.
- [166] D. Badre, J. Hoffman, J. W. Cooney, and M. D’esposito, “Hierarchical cognitive control deficits following damage to the human frontal lobe,” *Nature neuroscience*, vol. 12, no. 4, pp. 515–522, 2009.
- [167] J. Duncan, “The multiple-demand (md) system of the primate brain: mental programs for intelligent behaviour,” *Trends in cognitive sciences*, vol. 14, no. 4, pp. 172–179, 2010.
- [168] V. Mante, D. Sussillo, K. V. Shenoy, and W. T. Newsome, “Context-dependent computation by recurrent dynamics in prefrontal cortex,” *nature*, vol. 503, no. 7474, pp. 78–84, 2013.
- [169] M. G. Stokes, M. Kusunoki, N. Sigala, H. Nili, D. Gaffan, and J. Duncan, “Dynamic coding for cognitive control in prefrontal cortex,” *Neuron*, vol. 78, no. 2, pp. 364–375, 2013.
- [170] M. W. Cole, P. Laurent, and A. Stocco, “Rapid instructed task learning: A new window into the human brain’s unique capacity for flexible cognitive control,” *Cognitive, Affective, & Behavioral Neuroscience*, vol. 13, pp. 1–22, 2013.
- [171] P. Domenech and E. Koechlin, “Executive control and decision-making in the prefrontal cortex,” *Current opinion in behavioral sciences*, vol. 1, pp. 101–106, 2015.
- [172] M. C. Freund, J. A. Etzel, and T. S. Braver, “Neural coding of cognitive control: the representational similarity analysis approach,” *Trends in Cognitive Sciences*, vol. 25, no. 7, pp. 622–638, 2021.
- [173] D. Badre, A. Bhandari, H. Keglovits, and A. Kikumoto, “The dimensionality of neural representations for control,” *Current Opinion in Behavioral Sciences*, vol. 38, pp. 20–28, 2021.
- [174] A. W. MacDonald, J. D. Cohen, V. A. Stenger, and C. S. Carter, “Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control,” *Science*, vol. 288, no. 5472, pp. 1835–1838, 2000.
- [175] J. Duncan, H. Emslie, P. Williams, R. Johnson, and C. Freer, “Intelligence and the frontal lobe: The organization of goal-directed behavior,” *Cognitive psychology*, vol. 30, no. 3, pp. 257–303, 1996.
- [176] W. H. Alexander and J. W. Brown, “Medial prefrontal cortex as an action-outcome predictor,” *Nature neuroscience*, vol. 14, no. 10, pp. 1338–1344, 2011.

- [177] K. N. Ochsner and J. J. Gross, “The cognitive control of emotion,” *Trends in cognitive sciences*, vol. 9, no. 5, pp. 242–249, 2005.
- [178] D. Schiller, I. Levy, Y. Niv, J. E. LeDoux, and E. A. Phelps, “From fear to safety and back: reversal of fear in the human brain,” *Journal of Neuroscience*, vol. 28, no. 45, pp. 11517–11525, 2008.
- [179] N. Moneta, M. M. Garvert, H. R. Heekeren, and N. W. Schuck, “Task state representations in vmpfc mediate relevant and irrelevant value signals and their behavioral influence,” *Nature Communications*, vol. 14, no. 1, p. 3156, 2023.
- [180] D. Badre, R. A. Poldrack, E. J. Paré-Blagoev, R. Z. Insler, and A. D. Wagner, “Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex,” *Neuron*, vol. 47, no. 6, pp. 907–918, 2005.
- [181] D. Badre and A. D. Wagner, “Left ventrolateral prefrontal cortex and the cognitive control of memory,” *Neuropsychologia*, vol. 45, no. 13, pp. 2883–2901, 2007.
- [182] N. U. Dosenbach, D. A. Fair, F. M. Miezin, A. L. Cohen, K. K. Wenger, R. A. Dosenbach, M. D. Fox, A. Z. Snyder, J. L. Vincent, M. E. Raichle, *et al.*, “Distinct brain networks for adaptive and stable task control in humans,” *Proceedings of the National Academy of Sciences*, vol. 104, no. 26, pp. 11073–11078, 2007.
- [183] T. A. Jerde, E. P. Merriam, A. C. Riggall, J. H. Hedges, and C. E. Curtis, “Prioritized maps of space in human frontoparietal cortex,” *Journal of Neuroscience*, vol. 32, no. 48, pp. 17382–17390, 2012.
- [184] L. Cocchi, A. Zalesky, A. Fornito, and J. B. Mattingley, “Dynamic cooperation and competition between brain systems during cognitive control,” *Trends in cognitive sciences*, vol. 17, no. 10, pp. 493–501, 2013.
- [185] S. Sadaghiani and M. D’Esposito, “Functional characterization of the cingulo-opercular network in the maintenance of tonic alertness,” *Cerebral cortex*, vol. 25, no. 9, pp. 2763–2773, 2015.
- [186] T. Wu, X. Wang, Q. Wu, A. Spagna, J. Yang, C. Yuan, Y. Wu, Z. Gao, P. R. Hof, and J. Fan, “Anterior insular cortex is a bottleneck of cognitive control,” *NeuroImage*, vol. 195, pp. 490–504, 2019.
- [187] I. Molnar-Szakacs and L. Q. Uddin, “Anterior insula as a gatekeeper of executive control,” *Neuroscience & Biobehavioral Reviews*, vol. 139, p. 104736, 2022.

- [188] A. Kikumoto and U. Mayr, “Conjunctive representations that integrate stimuli, responses, and rules are critical for action selection,” *Proceedings of the National Academy of Sciences*, vol. 117, no. 19, pp. 10603–10608, 2020.
- [189] A. Kikumoto, A. Bhandari, K. Shibata, and D. Badre, “A transient high-dimensional geometry affords stable conjunctive subspaces for efficient action selection,” *Nature Communications*, vol. 15, no. 1, p. 8513, 2024.
- [190] S. F. de Kloet, B. Bruinsma, H. Terra, T. S. Heistek, E. M. Passchier, A. R. van den Berg, A. Luchicchi, R. Min, T. Pattij, and H. D. Mansvelder, “Bi-directional regulation of cognitive control by distinct prefrontal cortical output neurons to thalamus and striatum,” *Nature communications*, vol. 12, no. 1, pp. 1–15, 2021.
- [191] M. J. Frank, B. Loughry, and R. C. O'Reilly, “Interactions between frontal cortex and basal ganglia in working memory: a computational model,” *Cognitive, Affective, & Behavioral Neuroscience*, vol. 1, pp. 137–160, 2001.
- [192] X.-J. Wang, J. Tegnér, C. Constantinidis, and P. S. Goldman-Rakic, “Division of labor among distinct subtypes of inhibitory neurons in a cortical microcircuit of working memory,” *Proceedings of the National Academy of Sciences*, vol. 101, no. 5, pp. 1368–1373, 2004.
- [193] M. J. Frank, “Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making,” *Neural networks*, vol. 19, no. 8, pp. 1120–1136, 2006.
- [194] A. R. Aron and R. A. Poldrack, “Cortical and subcortical contributions to stop signal response inhibition: role of the subthalamic nucleus,” *Journal of Neuroscience*, vol. 26, no. 9, pp. 2424–2433, 2006.
- [195] A. R. Aron, “The neural basis of inhibition in cognitive control,” *The neuroscientist*, vol. 13, no. 3, pp. 214–228, 2007.
- [196] M. Medalla and H. Barbas, “Synapses with inhibitory neurons differentiate anterior cingulate from dorsolateral prefrontal pathways associated with cognitive control,” *Neuron*, vol. 61, no. 4, pp. 609–620, 2009.
- [197] A. R. Aron, T. W. Robbins, and R. A. Poldrack, “Inhibition and the right inferior frontal cortex: one decade on,” *Trends in cognitive sciences*, vol. 18, no. 4, pp. 177–185, 2014.

- [198] B. Casey, N. Tottenham, and J. Fossella, “Clinical, imaging, lesion, and genetic approaches toward a model of cognitive control,” *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, vol. 40, no. 3, pp. 237–254, 2002.
- [199] C. Liston, R. Watts, N. Tottenham, M. C. Davidson, S. Niogi, A. M. Ulug, and B. Casey, “Frontostriatal microstructure modulates efficient recruitment of cognitive control,” *Cerebral cortex*, vol. 16, no. 4, pp. 553–560, 2006.
- [200] V. M. Goghari and A. W. MacDonald III, “The neural basis of cognitive control: response selection and inhibition,” *Brain and cognition*, vol. 71, no. 2, pp. 72–83, 2009.
- [201] W. W. Seeley, V. Menon, A. F. Schatzberg, J. Keller, G. H. Glover, H. Kenna, A. L. Reiss, and M. D. Greicius, “Dissociable intrinsic connectivity networks for salience processing and executive control,” *Journal of neuroscience*, vol. 27, no. 9, pp. 2349–2356, 2007.
- [202] R. Cools, “Role of dopamine in the motivational and cognitive control of behavior,” *The Neuroscientist*, vol. 14, no. 4, pp. 381–395, 2008.
- [203] R. Cools and M. D’Esposito, “Inverted-u-shaped dopamine actions on human working memory and cognitive control,” *Biological psychiatry*, vol. 69, no. 12, pp. e113–e125, 2011.
- [204] E. C. Tolman, “Cognitive maps in rats and men.” *Psychological review*, vol. 55, no. 4, p. 189, 1948.
- [205] R. C. O’Reilly and M. J. Frank, “Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia,” *Neural computation*, vol. 18, no. 2, pp. 283–328, 2006.
- [206] R. C. O’Reilly, T. E. Hazy, and S. A. Herd, “The leabra cognitive architecture: How to play 20 principles with nature,” *The Oxford handbook of cognitive science*, vol. 91, pp. 91–116, 2016.
- [207] B. Noudoost, M. H. Chang, N. A. Steinmetz, and T. Moore, “Top-down control of visual attention,” *Current opinion in neurobiology*, vol. 20, no. 2, pp. 183–190, 2010.
- [208] L. Muller, F. Chavane, J. Reynolds, and T. J. Sejnowski, “Cortical travelling waves: mechanisms and computational principles,” *Nature Reviews Neuroscience*, vol. 19, no. 5, pp. 255–268, 2018.
- [209] A. R. Aron, D. M. Herz, P. Brown, B. U. Forstmann, and K. Zaghloul, “Frontosubthalamic circuits for control of action and cognition,” *Journal of Neuroscience*, vol. 36, no. 45, pp. 11489–11495, 2016.

- [210] M. Lundqvist, S. L. Brincat, J. Rose, M. R. Warden, T. J. Buschman, E. K. Miller, and P. Herman, “Working memory control dynamics follow principles of spatial computing,” *Nature Communications*, vol. 14, no. 1, p. 1429, 2023.
- [211] E. K. Miller, S. L. Brincat, and J. E. Roy, “Cognition is an emergent property,” *Current Opinion in Behavioral Sciences*, vol. 57, p. 101388, 2024.
- [212] L. Pessoa, “How do emotion and motivation direct executive control?,” *Trends in cognitive sciences*, vol. 13, no. 4, pp. 160–166, 2009.
- [213] T. A. Stalnaker, “Neuroscience: A bottom-up mechanism for cognitive control?,” *Current Biology*, vol. 34, no. 15, pp. R744–R746, 2024.
- [214] B. K. Leung, B. Chieng, S. Becchi, and B. W. Balleine, “A ventral pallidal-thalamocortical circuit mediates the cognitive control of instrumental action,” *Current Biology*, vol. 34, no. 15, pp. 3315–3326, 2024.
- [215] W. Hofmann, “Going beyond the individual level in self-control research,” *Nature Reviews Psychology*, vol. 3, no. 1, pp. 56–66, 2024.
- [216] N. Schneck, S. Herzog, J. Lu, A. Yttredahl, R. T. Ogden, H. Galfalvy, A. Burke, B. Stanley, J. J. Mann, and K. N. Ochsner, “The temporal dynamics of emotion regulation in subjects with major depression and healthy control subjects,” *Biological Psychiatry*, vol. 93, no. 3, pp. 260–267, 2023.
- [217] J. Rieck, J. Wrobel, A. R. Porras, K. McRae, and J. L. Gowin, “Neural signatures of emotion regulation,” *Scientific Reports*, vol. 14, no. 1, p. 1775, 2024.
- [218] S. Herzog, N. Schneck, H. Galfalvy, T. Hwei-Choo, M. Schmidt, C. A. Michel, M. E. Sublette, A. Burke, K. Ochsner, J. J. Mann, *et al.*, “A neural signature for reappraisal as an emotion regulation strategy: Relationship to stress-related suicidal ideation and negative affect in major depression,” *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, vol. 10, no. 1, pp. 94–102, 2025.
- [219] J. O’Neill, “The disciplinary society: from weber to foucault,” *British Journal of Sociology*, pp. 42–60, 1986.
- [220] R. Cooper and T. Shallice, “Contention scheduling and the control of routine activities,” *Cognitive neuropsychology*, vol. 17, no. 4, pp. 297–338, 2000.

- [221] M. Botvinick and D. C. Plaut, “Doing without schema hierarchies: a recurrent connectionist approach to normal and impaired routine sequential action.,” *Psychological review*, vol. 111, no. 2, p. 395, 2004.
- [222] R. P. Cooper and T. Shallice, “Hierarchical schemas and goals in the control of sequential behavior.,” 2006.
- [223] B. Prystawski, F. Mohnert, M. Tošić, and F. Lieder, “Resource-rational models of human goal pursuit,” *Topics in Cognitive Science*, vol. 14, no. 3, pp. 528–549, 2022.
- [224] P. Miller, C. D. Brody, R. Romo, and X.-J. Wang, “A recurrent network model of somatosensory parametric working memory in the prefrontal cortex,” *Cerebral Cortex*, vol. 13, no. 11, pp. 1208–1218, 2003.
- [225] T. E. Hazy, M. J. Frank, and R. C. O'Reilly, “Banishing the homunculus: making working memory work,” *Neuroscience*, vol. 139, no. 1, pp. 105–118, 2006.
- [226] F. van Ede, A. G. Board, and A. C. Nobre, “Goal-directed and stimulus-driven selection of internal representations,” *Proceedings of the National Academy of Sciences*, vol. 117, no. 39, pp. 24590–24598, 2020.
- [227] G. R. Yang, M. R. Joglekar, H. F. Song, W. T. Newsome, and X.-J. Wang, “Task representations in neural networks trained to perform many cognitive tasks,” *Nature neuroscience*, vol. 22, no. 2, pp. 297–306, 2019.
- [228] R. N. Spreng, E. DuPre, D. Selarka, J. Garcia, S. Gojkovic, J. Mildner, W.-M. Luh, and G. R. Turner, “Goal-congruent default network activity facilitates cognitive control,” *Journal of Neuroscience*, vol. 34, no. 42, pp. 14108–14114, 2014.
- [229] C. Kidd, H. Palmeri, and R. N. Aslin, “Rational snacking: Young children's decision-making on the marshmallow task is moderated by beliefs about environmental reliability,” *Cognition*, vol. 126, no. 1, pp. 109–114, 2013.
- [230] E. Sutlief, C. Walters, T. Marton, and M. G. H. Shuler, “The value of initiating a pursuit in temporal decision-making,” *Elife*, vol. 13, p. RP99957, 2025.
- [231] W. Mischel and E. B. Ebbesen, “Attention in delay of gratification.,” *Journal of personality and social psychology*, vol. 16, no. 2, p. 329, 1970.

- [232] J. Heffner, J.-Y. Son, and O. FeldmanHall, “Emotion prediction errors guide socially adaptive behaviour,” *Nature human behaviour*, vol. 5, no. 10, pp. 1391–1401, 2021.
- [233] R. Schurr, D. Reznik, H. Hillman, R. Bhui, and S. J. Gershman, “Dynamic computational phenotyping of human cognition,” *Nature Human Behaviour*, vol. 8, no. 5, pp. 917–931, 2024.
- [234] A. Shenhav, “The affective gradient hypothesis: an affect-centered account of motivated behavior,” *Trends in Cognitive Sciences*, 2024.
- [235] R. M. Ryan, E. L. Deci, and W. S. Grolnick, “Autonomy, relatedness, and the self: Their relation to development and psychopathology.” 1995.
- [236] E. L. Deci and R. M. Ryan, *Intrinsic motivation and self-determination in human behavior*. Springer Science & Business Media, 2013.
- [237] D. A. Robson, M. S. Allen, and S. J. Howard, “Self-regulation in childhood as a predictor of future outcomes: A meta-analytic review.”, *Psychological bulletin*, vol. 146, no. 4, p. 324, 2020.
- [238] J. Chu, J. B. Tenenbaum, and L. E. Schulz, “In praise of folly: flexible goals and human cognition,” *Trends in Cognitive Sciences*, vol. 28, no. 7, pp. 628–642, 2024.
- [239] Y. Munakata and L. E. Michaelson, “Executive functions in social context: Implications for conceptualizing, measuring, and supporting developmental trajectories,” *Annual Review of Developmental Psychology*, vol. 3, no. 1, pp. 139–163, 2021.
- [240] M. Inzlicht and E. Berkman, “Six questions for the resource model of control (and some answers),” *Social and Personality Psychology Compass*, vol. 9, no. 10, pp. 511–524, 2015.
- [241] D. Randles, I. Harlow, and M. Inzlicht, “A pre-registered naturalistic observation of within domain mental fatigue and domain-general depletion of self-control,” *PLoS one*, vol. 12, no. 9, p. e0182980, 2017.
- [242] A. Z. Enkavi, I. W. Eisenberg, P. G. Bissett, G. L. Mazza, D. P. MacKinnon, L. A. Marsch, and R. A. Poldrack, “Large-scale analysis of test-retest reliabilities of self-regulation measures,” *Proceedings of the National Academy of Sciences*, vol. 116, no. 12, pp. 5472–5477, 2019.
- [243] A. G. Collins and M. J. Frank, “Cognitive control over learning: creating, clustering, and generalizing task-set structure.,” *Psychological review*, vol. 120, no. 1, p. 190, 2013.

- [244] A. L. Duckworth, T. S. Gendler, and J. J. Gross, “Situational strategies for self-control,” *Perspectives on Psychological Science*, vol. 11, no. 1, pp. 35–55, 2016.
- [245] A. L. Duckworth, K. L. Milkman, and D. Laibson, “Beyond willpower: Strategies for reducing failures of self-control,” *Psychological Science in the Public Interest*, vol. 19, no. 3, pp. 102–129, 2018.
- [246] N. P. Friedman, A. Miyake, R. P. Corley, S. E. Young, J. C. DeFries, and J. K. Hewitt, “Not all executive functions are related to intelligence,” *Psychological science*, vol. 17, no. 2, pp. 172–179, 2006.
- [247] M. Inzlicht and M. Fries, “The past, present, and future of ego depletion,” *Social Psychology*, 2019.
- [248] L. G. Smithers, A. C. Sawyer, C. R. Chittleborough, N. M. Davies, G. Davey Smith, and J. W. Lynch, “A systematic review and meta-analysis of effects of early life non-cognitive skills on academic, psychosocial, cognitive and health outcomes,” *Nature human behaviour*, vol. 2, no. 11, pp. 867–880, 2018.
- [249] A. Baddeley, “Fractionating the central executive,” *Principles of frontal lobe function*, pp. 246–260, 2002.
- [250] P. Drucker, *The practice of management*. Routledge, 1954.
- [251] H. A. Simon, *Administrative behavior*. Simon and Schuster, 2013.
- [252] F. W. Taylor, *The principles of scientific management*. Harper & brothers, 1919.
- [253] G. T. Doran, “There’s a s.m.a.r.t. way to write management’s goals and objectives,” *Management Review*, vol. 70, no. 11, pp. 35–36, 1981.
- [254] L. N. Ross, “Explanation in biology,” *Elements in the Philosophy of Biology*, 2025.
- [255] D. E. Kirk, *Optimal control theory: an introduction*. Courier Corporation, 2004.
- [256] R. S. Marken and W. Mansell, “Perceptual control as a unifying concept in psychology,” *Review of General Psychology*, vol. 17, no. 2, pp. 190–195, 2013.
- [257] E. Todorov and M. I. Jordan, “Optimal feedback control as a theory of motor coordination,” *Nature neuroscience*, vol. 5, no. 11, pp. 1226–1235, 2002.
- [258] R. Shadmehr, H. J. Huang, and A. A. Ahmed, “Effort, reward, and vigor in decision-making and motor control,” *Current biology: CB*, vol. 26, no. 14, p. 1929, 2016.

- [259] B. H. Scheid, A. Ashourvan, J. Stiso, K. A. Davis, F. Mikhail, F. Pasqualetti, B. Litt, and D. S. Bassett, “Time-evolving controllability of effective connectivity networks during seizure progression,” *Proceedings of the National Academy of Sciences*, vol. 118, no. 5, p. e2006436118, 2021.
- [260] T. C. Rouse, A. M. Ni, C. Huang, and M. R. Cohen, “Topological insights into the neural basis of flexible behavior,” *Proceedings of the National Academy of Sciences*, vol. 120, no. 24, p. e2219557120, 2023.
- [261] A. Avena-Koenigsberger, B. Misic, and O. Sporns, “Communication dynamics in complex brain networks,” *Nature reviews neuroscience*, vol. 19, no. 1, pp. 17–33, 2018.
- [262] D. Zhou, C. W. Lynn, Z. Cui, R. Ciric, G. L. Baum, T. M. Moore, D. R. Roalf, J. A. Detre, R. C. Gur, R. E. Gur, *et al.*, “Efficient coding in the economics of human brain connectomics,” *Network Neuroscience*, vol. 6, no. 1, pp. 234–274, 2022.
- [263] C. Seguin, O. Sporns, and A. Zalesky, “Brain network communication: concepts, models and applications,” *Nature reviews neuroscience*, vol. 24, no. 9, pp. 557–574, 2023.
- [264] D. Zhou, Y. Kang, D. Cosme, M. Jovanova, X. He, A. Mahadevan, J. Ahn, O. Stanoi, J. K. Brynildsen, N. Cooper, *et al.*, “Mindful attention promotes control of brain network dynamics for self-regulation and discontinues the past from the present,” *Proceedings of the National Academy of Sciences*, vol. 120, no. 2, p. e2201074119, 2023.
- [265] D. S. Bassett, D. L. Greenfield, A. Meyer-Lindenberg, D. R. Weinberger, S. W. Moore, and E. T. Bullmore, “Efficient physical embedding of topologically complex information processing networks in brains and computer circuits,” *PLoS computational biology*, vol. 6, no. 4, p. e1000748, 2010.
- [266] D. S. Bassett and E. Bullmore, “Small-world brain networks,” *The neuroscientist*, vol. 12, no. 6, pp. 512–523, 2006.
- [267] P. Sterling and S. Laughlin, *Principles of neural design*. MIT press, 2015.
- [268] W. Ashby, *Design for a brain: The origin of adaptive behaviour*. Springer Science & Business Media, 1952.
- [269] P. Sterling, “Allostasis: a model of predictive regulation,” *Physiology & behavior*, vol. 106, no. 1, pp. 5–15, 2012.

- [270] D. Zhou, I. Tseytlin, T. Satterthwaite, and D. S. Bassett, “Predictive coding from compression, control, and recurrence in human brain networks,” in *2023 Conference on Cognitive Computational Neuroscience*, 2023.
- [271] J. E. Theriault, Y. Katsumi, H. M. Reimann, J. Zhang, P. Deming, B. C. Dickerson, K. S. Quigley, and L. F. Barrett, “It’s not the thought that counts: Allostasis at the core of brain function,” *Neuron*, 2025.
- [272] S. Z. Mehrhof, H. Fleming, and C. L. Nord, “An interoceptive model of energy allostasis linking metabolic and mental health,” *Science Advances*, vol. 11, no. 39, p. eady4356, 2025.
- [273] T. Yarkoni, R. A. Poldrack, T. E. Nichols, D. C. Van Essen, and T. D. Wager, “Large-scale automated synthesis of human functional neuroimaging data,” *Nature methods*, vol. 8, no. 8, pp. 665–670, 2011.
- [274] R. F. Betzel, S. Gu, J. D. Medaglia, F. Pasqualetti, and D. S. Bassett, “Optimally controlling the human connectome: the role of network topology,” *Scientific reports*, vol. 6, no. 1, p. 30770, 2016.
- [275] C. Gratton, H. Sun, and S. E. Petersen, “Control networks and hubs,” *Psychophysiology*, vol. 55, no. 3, p. e13032, 2018.
- [276] S. P. Patankar, J. Z. Kim, F. Pasqualetti, and D. S. Bassett, “Path-dependent connectivity, not modularity, consistently predicts controllability of structural brain networks,” *Network Neuroscience*, vol. 4, no. 4, pp. 1091–1121, 2020.
- [277] Z. Cui, J. Stiso, G. L. Baum, J. Z. Kim, D. R. Roalf, R. F. Betzel, S. Gu, Z. Lu, C. H. Xia, X. He, *et al.*, “Optimization of energy state transition trajectory supports the development of executive function during youth,” *elife*, vol. 9, p. e53060, 2020.
- [278] D. Zhou, J. Ahn, D. M. Lydon-Staley, E. B. Falk, D. S. Bassett, and A. M. Ruscio, “Neural dynamics underlying the persistence of perseverative thought in depression and anxiety,” *bioRxiv*, pp. 2025–07, 2025.
- [279] S. Gu, F. Pasqualetti, M. Cieslak, Q. K. Telesford, A. B. Yu, A. E. Kahn, J. D. Medaglia, J. M. Vettel, M. B. Miller, S. T. Grafton, *et al.*, “Controllability of structural brain networks,” *Nature communications*, vol. 6, no. 1, p. 8414, 2015.
- [280] R. B. Lopez, W. Hofmann, D. D. Wagner, W. M. Kelley, and T. F. Heatherton, “Neural predictors of giving in to temptation in daily life,” *Psychological science*, vol. 25, no. 7, pp. 1337–1344, 2014.

- [281] Y.-C. Chiu, J. Jiang, and T. Egner, “The caudate nucleus mediates learning of stimulus–control state associations,” *Journal of Neuroscience*, vol. 37, no. 4, pp. 1028–1038, 2017.
- [282] M. Lengyel and P. Dayan, “Hippocampal contributions to control: the third way,” *Advances in neural information processing systems*, vol. 20, 2007.
- [283] A. Chung, C. Jou, A. Grau-Perales, E. R. Levy, D. Dvorak, N. Hussain, and A. A. Fenton, “Cognitive control persistently enhances hippocampal information processing,” *Nature*, vol. 600, no. 7889, pp. 484–488, 2021.
- [284] M. Agrawal, M. G. Mattar, J. D. Cohen, and N. D. Daw, “The temporal dynamics of opportunity costs: A normative account of cognitive fatigue and boredom.,” *Psychological review*, vol. 129, no. 3, p. 564, 2022.
- [285] E. H. Park, K. C. O. Sparks, G. Grubbs, D. Taborga, K. Nicholas, A. S. Ahmed, N. Ruiz-Pérez, N. Kim, S. Segura-Carrillo, and A. A. Fenton, “Cognitive control of behavior and hippocampal information processing without medial prefrontal cortex,” *eLife*, vol. 13, 2025.
- [286] J. M. Fine and B. Y. Hayden, “The whole prefrontal cortex is premotor cortex,” *Philosophical Transactions of the Royal Society B*, vol. 377, no. 1844, p. 20200524, 2022.
- [287] C. D. Salzman and S. Fusi, “Emotion, cognition, and mental state representation in amygdala and prefrontal cortex,” *Annual review of neuroscience*, vol. 33, no. 1, pp. 173–202, 2010.
- [288] M. Inzlicht, B. D. Bartholow, and J. B. Hirsh, “Emotional foundations of cognitive control,” *Trends in cognitive sciences*, vol. 19, no. 3, pp. 126–132, 2015.
- [289] M.-M. Mesulam, “From sensation to cognition.,” *Brain: a journal of neurology*, vol. 121, no. 6, pp. 1013–1052, 1998.
- [290] M. P. Van Den Heuvel, R. S. Kahn, J. Goñi, and O. Sporns, “High-cost, high-capacity backbone for global brain communication,” *Proceedings of the National Academy of Sciences*, vol. 109, no. 28, pp. 11372–11377, 2012.
- [291] C. Gratton, T. O. Laumann, E. M. Gordon, B. Adeyemo, and S. E. Petersen, “Evidence for two independent factors that modify brain networks to meet task goals,” *Cell reports*, vol. 17, no. 5, pp. 1276–1288, 2016.

- [292] D. Zhou, J. Z. Kim, A. R. Pines, V. J. Sydnor, D. R. Roalf, J. A. Detre, R. C. Gur, R. E. Gur, T. D. Satterthwaite, and D. S. Bassett, “Compression supports low-dimensional representations of behavior across neural circuits,” *bioRxiv*, pp. 2022–11, 2022.
- [293] R. T. Canolty and R. T. Knight, “The functional role of cross-frequency coupling,” *Trends in cognitive sciences*, vol. 14, no. 11, pp. 506–515, 2010.
- [294] J. Z. Kim, J. M. Soffer, A. E. Kahn, J. M. Vettel, F. Pasqualetti, and D. S. Bassett, “Role of graph architecture in controlling dynamical networks with applications to neural systems,” *Nature physics*, vol. 14, no. 1, pp. 91–98, 2018.
- [295] E. Tang, C. Giusti, G. L. Baum, S. Gu, E. Pollock, A. E. Kahn, D. R. Roalf, T. M. Moore, K. Ruparel, R. C. Gur, *et al.*, “Developmental increases in white matter network controllability support a growing diversity of brain dynamics,” *Nature communications*, vol. 8, no. 1, p. 1252, 2017.
- [296] J. Lim, I. Mitrai, P. Daoutidis, and C. Stamoulis, “The adolescent functional connectome is dynamically controlled by a sparse core of cognitive and topological hubs,” *bioRxiv*, pp. 2025–03, 2025.
- [297] U. Braun, A. Harneit, G. Pergola, T. Menara, A. Schäfer, R. F. Betzel, Z. Zang, J. I. Schweiger, X. Zhang, K. Schwarz, *et al.*, “Brain network dynamics during working memory are modulated by dopamine and diminished in schizophrenia,” *Nature Communications*, vol. 12, no. 1, pp. 1–11, 2021.
- [298] K. P. Szymula, F. Pasqualetti, A. M. Graybiel, T. M. Desrochers, and D. S. Bassett, “Habit learning supported by efficiently controlled network dynamics in naive macaque monkeys,” *arXiv*, vol. 2006, no. 14565, 2020.
- [299] E. J. Cornblath, A. Ashourvan, J. Z. Kim, R. F. Betzel, R. Ceric, A. Adebimpe, G. L. Baum, X. He, K. Ruparel, T. M. Moore, *et al.*, “Temporal sequences of brain activity at rest are constrained by white matter structure and modulated by cognitive demands,” *Communications biology*, vol. 3, no. 1, p. 261, 2020.
- [300] A. I. Luppi, S. P. Singleton, J. Y. Hansen, K. W. Jamison, D. Bzdok, A. Kuceyeski, R. F. Betzel, and B. Misic, “Contributions of network structure, chemoarchitecture and diagnostic categories to transitions between cognitive topographies,” *Nature Biomedical Engineering*, vol. 8, no. 9, pp. 1142–1161, 2024.

- [301] E. J. Cornblath, E. Tang, G. L. Baum, T. M. Moore, A. Adebimpe, D. R. Roalf, R. C. Gur, R. E. Gur, F. Pasqualetti, T. D. Satterthwaite, *et al.*, “Sex differences in network controllability as a predictor of executive function in youth,” *NeuroImage*, vol. 188, pp. 122–134, 2019.
- [302] W. Cai, S. Ryali, R. Pasumarthy, V. Talasila, and V. Menon, “Dynamic causal brain circuits during working memory and their functional controllability,” *Nature communications*, vol. 12, no. 1, p. 3314, 2021.
- [303] L. Parkes, J. Z. Kim, J. Stiso, M. E. Calkins, M. Cieslak, R. E. Gur, R. C. Gur, T. M. Moore, M. Ouellet, D. R. Roalf, *et al.*, “Asymmetric signaling across the hierarchy of cytoarchitecture within the human connectome,” *Science Advances*, vol. 8, no. 50, p. eadd2185, 2022.
- [304] L. Weninger, P. Srivastava, D. Zhou, J. Z. Kim, E. J. Cornblath, M. A. Bertolero, U. Habel, D. Merhof, and D. S. Bassett, “Information content of brain states is explained by structural constraints on state energetics,” *Physical Review E*, vol. 106, no. 1, p. 014401, 2022.
- [305] E. G. Ceballos, A. I. Luppi, G. Castrillon, M. Saggar, B. Misic, and V. Riedl, “The control costs of human brain dynamics,” *Network Neuroscience*, vol. 9, no. 1, p. 77, 2025.
- [306] S. P. Singleton, A. I. Luppi, R. L. Carhart-Harris, J. Cruzat, L. Roseman, G. Deco, M. L. Kringelbach, E. A. Stamatakis, and A. Kuceyeski, “Lsd flattens the brain’s energy landscape: evidence from receptor-informed network control theory,” *bioRxiv*, 2021.
- [307] A. S. Mahadevan, E. J. Cornblath, D. M. Lydon-Staley, D. Zhou, L. Parkes, B. Larsen, A. Adebimpe, A. E. Kahn, R. C. Gur, R. E. Gur, *et al.*, “Alprazolam modulates persistence energy during emotion processing in first-degree relatives of individuals with schizophrenia: a network control study,” *bioRxiv*, 2021.
- [308] X. He, L. Caciagli, L. M. Parkes, J. Stiso, T. M. Karrer, J. Z. Kim, Z. Lu, T. Menara, F. Pasqualetti, M. R. Sperling, *et al.*, “Pathological and metabolic underpinnings of energetic inefficiency in temporal lobe epilepsy,” *bioRxiv*, 2021.
- [309] J. M. Shine, M. Breakspear, P. T. Bell, K. A. Ehgoetz Martens, R. Shine, O. Koyejo, O. Sporns, and R. A. Poldrack, “Human cognition involves the dynamic integration of neural activity and neuromodulatory systems,” *Nature neuroscience*, vol. 22, no. 2, pp. 289–296, 2019.

- [310] R. Cools and A. F. Arnsten, “Neuromodulation of prefrontal cortex cognitive function in primates: the powerful roles of monoamines and acetylcholine,” *Neuropsychopharmacology*, vol. 47, no. 1, pp. 309–328, 2022.
- [311] N. Wiener, *Cybernetics or Control and Communication in the Animal and the Machine*. MIT press, 2019.
- [312] C. B. Holroyd, J. J. Ribas-Fernandes, D. Shahnazian, M. Silvetti, and T. Verguts, “Human midcingulate cortex encodes distributed representations of task progress,” *Proceedings of the National Academy of Sciences*, vol. 115, no. 25, pp. 6398–6403, 2018.
- [313] P. Dayan, “How to set the switches on this thing,” *Current opinion in neurobiology*, vol. 22, no. 6, pp. 1068–1074, 2012.
- [314] A. Gomez-Marin and A. A. Ghazanfar, “The life of behavior,” *Neuron*, vol. 104, no. 1, pp. 25–36, 2019.
- [315] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, “Control centrality and hierarchical structure in complex networks,” 2012.
- [316] T. R. Henry, D. J. Robinaugh, and E. I. Fried, “On the control of psychological networks,” *Psychometrika*, vol. 87, no. 1, pp. 188–213, 2022.
- [317] J. E. Stocker, G. Koppe, H. Reich, S. Heshmati, S. Kittel-Schneider, S. G. Hofmann, T. Hahn, H. L. van der Maas, L. Waldorp, and H. Jamalabadi, “Formalizing psychological interventions through network control theory,” *Scientific Reports*, vol. 13, no. 1, p. 13830, 2023.
- [318] C.-T. Chen, *Linear system theory and design*. Saunders college publishing, 1984.
- [319] N. J. Cowan, E. J. Chastain, D. A. Vilhena, J. S. Freudenberg, and C. T. Bergstrom, “Nodal dynamics, not degree distributions, determine the structural controllability of complex networks,” *PloS one*, vol. 7, no. 6, p. e38398, 2012.
- [320] C. O. Becker, S. Pequito, G. J. Pappas, and V. M. Preciado, “Network design for controllability metrics,” *IEEE Transactions on Control of Network Systems*, vol. 7, no. 3, pp. 1404–1415, 2020.
- [321] W. R. Ashby, “An introduction to cybernetics,” 1956.
- [322] M. Rubinov, “Circular and unified analysis in network neuroscience,” 2022.

- [323] S. F. Muldoon, F. Pasqualetti, S. Gu, M. Cieslak, S. T. Grafton, J. M. Vettel, and D. S. Bassett, “Stimulation-based control of dynamic brain networks,” *PLoS computational biology*, vol. 12, no. 9, p. e1005076, 2016.
- [324] E. Nozari, J. Stiso, L. Caciagli, E. J. Cornblath, X. He, M. A. Bertolero, A. S. Mahadevan, G. J. Pappas, and D. S. Bassett, “Is the brain macroscopically linear? a system identification of resting state dynamics,” *bioRxiv*, 2020.
- [325] A. T. Buss and J. P. Spencer, “The emergent executive: A dynamic field theory of the development of executive function,” *Monographs of the Society for Research in Child Development*, vol. 79, no. 2, p. vii, 2014.
- [326] L. E. Suárez, A. Mihalik, F. Milisav, K. Marshall, M. Li, P. E. Vértes, G. Lajoie, and B. Misic, “Connectome-based reservoir computing with the conn2res toolbox,” *Nature Communications*, vol. 15, no. 1, p. 656, 2024.
- [327] S. L. Brunton, J. L. Proctor, and J. N. Kutz, “Discovering governing equations from data by sparse identification of nonlinear dynamical systems,” *Proceedings of the national academy of sciences*, vol. 113, no. 15, pp. 3932–3937, 2016.
- [328] D. Durstewitz, G. Koppe, and M. I. Thurm, “Reconstructing computational system dynamics from neural data with recurrent neural networks,” *Nature Reviews Neuroscience*, vol. 24, no. 11, pp. 693–710, 2023.
- [329] M. Brenner, E. Weber, G. Koppe, and D. Durstewitz, “Learning interpretable hierarchical dynamical systems models from time series data,” *arXiv preprint arXiv:2410.04814*, 2024.
- [330] G. Yan, P. E. Vértes, E. K. Towlson, Y. L. Chew, D. S. Walker, W. R. Schafer, and A.-L. Barabási, “Network control principles predict neuron function in the *caenorhabditis elegans* connectome,” *Nature*, vol. 550, no. 7677, pp. 519–523, 2017.
- [331] J. K. Brynildsen, K. D. Mace, E. J. Cornblath, C. Weidler, F. Pasqualetti, D. S. Bassett, and J. A. Blendy, “Gene coexpression patterns predict opiate-induced brain-state transitions,” *Proceedings of the National Academy of Sciences*, vol. 117, no. 32, pp. 19556–19565, 2020.

- [332] J. Stiso, A. N. Khambhati, T. Menara, A. E. Kahn, J. M. Stein, S. R. Das, R. Gorniak, J. Tracy, B. Litt, K. A. Davis, *et al.*, “White matter network architecture guides direct electrical stimulation through optimal state transitions,” *Cell reports*, vol. 28, no. 10, pp. 2554–2566, 2019.
- [333] L. Parkes, T. M. Moore, M. E. Calkins, M. Cieslak, D. R. Roalf, D. H. Wolf, R. C. Gur, R. E. Gur, T. D. Satterthwaite, and D. S. Bassett, “Network controllability in transmodal cortex predicts positive psychosis spectrum symptoms,” *Biological Psychiatry*, vol. 90, no. 6, pp. 409–418, 2021.
- [334] D. Zöller, C. Sandini, M. Schaer, S. Eliez, D. S. Bassett, and D. Van De Ville, “Structural control energy of resting-state functional brain states reveals less cost-effective brain dynamics in psychosis vulnerability,” *Human brain mapping*, vol. 42, no. 7, pp. 2181–2200, 2021.
- [335] C. Dimulescu, S. Gareayaghi, F. Kamp, S. Fromm, K. Obermayer, and C. Metzner, “Structural differences between healthy subjects and patients with schizophrenia or schizoaffective disorder: A graph and control theoretical perspective,” *Frontiers in Psychiatry*, p. 991, 2021.
- [336] Y. N. Kenett, R. E. Beaty, and J. D. Medaglia, “A computational network control theory analysis of depression symptoms,” *Personality neuroscience*, vol. 1, p. e16, 2018.
- [337] T. Hahn, H. Jamalabadi, D. Emden, J. Goltermann, J. Ernsting, N. R. Winter, L. Fisch, R. Leenings, K. Sarink, V. Holstein, *et al.*, “A network control theory approach to longitudinal symptom dynamics in major depressive disorder,” *arXiv preprint arXiv:2107.10178*, 2021.
- [338] T. Hahn, N. R. Winter, J. Ernsting, M. Gruber, M. J. Mauritz, L. Fisch, R. Leenings, K. Sarink, J. Blanke, V. Holstein, *et al.*, “Genetic, individual, and familial risk correlates of brain network controllability in major depressive disorder,” *Molecular psychiatry*, vol. 28, no. 3, pp. 1057–1063, 2023.
- [339] T. R. Henry, N. D. Fogelman, T. Nugiel, and J. R. Cohen, “Effect of methylphenidate on functional controllability: a preliminary study in medication-naïve children with adhd,” *Translational Psychiatry*, vol. 12, no. 1, p. 518, 2022.
- [340] J. Jeganathan, A. Perry, D. S. Bassett, G. Roberts, P. B. Mitchell, and M. Breakspear, “Fronto-limbic dysconnectivity leads to impaired brain network controllability in young people with bipolar disorder and those at high genetic risk,” *NeuroImage: Clinical*, vol. 19, pp. 71–81, 2018.

- [341] B. Wang, S. Zhang, X. Yu, Y. Niu, J. Niu, D. Li, S. Zhang, J. Xiang, T. Yan, J. Yang, *et al.*, “Alterations in white matter network dynamics in patients with schizophrenia and bipolar disorder,” *Human Brain Mapping*, vol. 43, no. 13, pp. 3909–3922, 2022.
- [342] A. Chari, K. K. Seunarine, X. He, M. M. Tisdall, C. A. Clark, D. S. Bassett, R. C. Scott, and R. E. Rosch, “Drug-resistant focal epilepsy in children is associated with increased modal controllability of the whole brain and epileptogenic regions,” *Communications Biology*, vol. 5, no. 1, p. 394, 2022.
- [343] L. Beynel, L. Deng, C. A. Crowell, M. Dannhauer, H. Palmer, S. Hilbig, A. V. Peterchev, B. Luber, S. H. Lisanby, R. Cabeza, *et al.*, “Structural controllability predicts functional patterns and brain stimulation benefits associated with working memory,” *Journal of Neuroscience*, vol. 40, no. 35, pp. 6770–6778, 2020.
- [344] F. Fang, B. Godlewska, R. Y. Cho, S. I. Savitz, S. Selvaraj, and Y. Zhang, “Personalizing repetitive transcranial magnetic stimulation for precision depression treatment based on functional brain network controllability and optimal control analysis,” *NeuroImage*, vol. 260, p. 119465, 2022.
- [345] J. Wilmskroetter, X. He, L. Caciagli, J. H. Jensen, B. Marebwa, K. A. Davis, J. Fridriksson, A. Basilakos, L. P. Johnson, C. Rorden, *et al.*, “Language recovery after brain injury: A structural network control theory study,” *Journal of Neuroscience*, vol. 42, no. 4, pp. 657–669, 2022.
- [346] T. S. Braver, M. K. Krug, K. S. Chiew, W. Kool, J. A. Westbrook, N. J. Clement, R. A. Adcock, D. M. Barch, M. M. Botvinick, C. S. Carver, *et al.*, “Mechanisms of motivation–cognition interaction: challenges and opportunities,” *Cognitive, Affective, & Behavioral Neuroscience*, vol. 14, no. 2, pp. 443–472, 2014.
- [347] J. M. Bugg, L. L. Jacoby, and J. P. Toth, “Multiple levels of control in the stroop task,” *Memory & cognition*, vol. 36, no. 8, pp. 1484–1494, 2008.
- [348] P.-S. Lee and D. K. Sewell, “A revised diffusion model for conflict tasks,” *Psychonomic Bulletin & Review*, vol. 31, no. 1, pp. 1–31, 2024.
- [349] A. Dey and J. M. Bugg, “The timescale of control: A meta-control property that generalizes across tasks but varies between types of control,” *Cognitive, Affective, & Behavioral Neuroscience*, vol. 21, pp. 472–489, 2021.

- [350] J. Z. Kim, R. F. Betzel, A. Beyh, A. Howell, A. Kuceyeski, B. Larsen, C. Seguin, X.-H. Zhang, A. Holmes, and L. Parkes, “Inferring macroscopic intrinsic neural timescales using optimal control theory,” *bioRxiv*, pp. 2025–04, 2025.
- [351] M. E. Hasselmo, “A model of prefrontal cortical mechanisms for goal-directed behavior,” *Journal of cognitive neuroscience*, vol. 17, no. 7, pp. 1115–1129, 2005.
- [352] J. F. Thayer, A. L. Hansen, E. Saus-Rose, and B. H. Johnsen, “Heart rate variability, prefrontal neural function, and cognitive performance: the neurovisceral integration perspective on self-regulation, adaptation, and health,” *Annals of behavioral medicine*, vol. 37, no. 2, pp. 141–153, 2009.
- [353] L. F. Barrett and W. K. Simmons, “Interoceptive predictions in the brain,” *Nature reviews neuroscience*, vol. 16, no. 7, pp. 419–429, 2015.
- [354] C. Baiano, G. Santangelo, V. P. Senese, G. Di Mauro, G. Lauro, M. Piacenti, and M. Conson, “Linking perception of bodily states and cognitive control: the role of interoception in impulsive behaviour,” *Experimental Brain Research*, vol. 239, no. 3, pp. 857–865, 2021.
- [355] E. Sennesh, J. Theriault, D. Brooks, J.-W. van de Meent, L. F. Barrett, and K. S. Quigley, “Interoception as modeling, allostasis as control,” *Biological psychology*, vol. 167, p. 108242, 2022.
- [356] A. Kok, “Cognitive control, motivation and fatigue: A cognitive neuroscience perspective,” *Brain and Cognition*, vol. 160, p. 105880, 2022.
- [357] R. N. Spreng and C. L. Grady, “Patterns of brain activity supporting autobiographical memory, prospection, and theory of mind, and their relationship to the default mode network,” *Journal of cognitive neuroscience*, vol. 22, no. 6, pp. 1112–1123, 2010.
- [358] L. Koban, P. J. Gianaros, H. Kober, and T. D. Wager, “The self in context: brain systems linking mental and physical health,” *Nature Reviews Neuroscience*, vol. 22, no. 5, pp. 309–322, 2021.
- [359] V. Menon, “20 years of the default mode network: A review and synthesis,” *Neuron*, vol. 111, no. 16, pp. 2469–2487, 2023.
- [360] P. Srivastava, P. J. Mucha, E. Falk, F. Pasqualetti, and D. S. Bassett, “Structural underpinnings of control in multiplex networks,” *arXiv preprint arXiv:2103.08757*, 2021.

- [361] E. B. Falk and D. S. Bassett, “Brain and social networks: fundamental building blocks of human experience,” *Trends in cognitive sciences*, vol. 21, no. 9, pp. 674–690, 2017.
- [362] R. C. Wilson, Y. K. Takahashi, G. Schoenbaum, and Y. Niv, “Orbitofrontal cortex as a cognitive map of task space,” *Neuron*, vol. 81, no. 2, pp. 267–279, 2014.
- [363] N. W. Schuck, M. B. Cai, R. C. Wilson, and Y. Niv, “Human orbitofrontal cortex represents a cognitive map of state space,” *Neuron*, vol. 91, no. 6, pp. 1402–1412, 2016.
- [364] M. L. Mack, A. R. Preston, and B. C. Love, “Ventromedial prefrontal cortex compression during concept learning,” *Nature communications*, vol. 11, no. 1, p. 46, 2020.
- [365] E. C. Tolman, *Purposive behavior in animals and men*. Univ of California Press, 1951.
- [366] E. Kelemen and A. A. Fenton, “Coordinating different representations in the hippocampus,” *Neurobiology of learning and memory*, vol. 129, pp. 50–59, 2016.
- [367] J. L. Bellmund, P. Gärdenfors, E. I. Moser, and C. F. Doeller, “Navigating cognition: Spatial codes for human thinking,” *Science*, vol. 362, no. 6415, p. eaat6766, 2018.
- [368] I. K. Brunec, B. Bellana, J. D. Ozubko, V. Man, J. Robin, Z.-X. Liu, C. Grady, R. S. Rosenbaum, G. Winocur, M. D. Barense, *et al.*, “Multiple scales of representation along the hippocampal antero-posterior axis in humans,” *Current biology*, vol. 28, no. 13, pp. 2129–2135, 2018.
- [369] R. W. White, “Motivation reconsidered: the concept of competence.,” *Psychological review*, vol. 66, no. 5, p. 297, 1959.
- [370] A. Bandura, “Self-efficacy: toward a unifying theory of behavioral change.,” *Psychological review*, vol. 84, no. 2, p. 191, 1977.
- [371] E. T. Higgins, *Beyond pleasure and pain: How motivation works*. Oxford University Press, 2011.
- [372] G. Molinaro and A. G. Collins, “A goal-centric outlook on learning,” *Trends in Cognitive Sciences*, 2023.
- [373] G. Pistilli, A. Leidinger, Y. Jernite, A. Kasirzadeh, A. S. Luccioni, and M. Mitchell, “Civics: Building a dataset for examining culturally-informed values in large language models,” in *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, vol. 7, pp. 1132–1144, 2024.

- [374] S. Bhatia, S. T. van Baal, F. Wang, and L. Walasek, “Computational analysis of 100 k choice dilemmas: Decision attributes, trade-off structures, and model-based prediction,” *Proceedings of the National Academy of Sciences*, vol. 122, no. 17, p. e2406489122, 2025.
- [375] D. Tamir, F. Corbani, J. Mildner, and D. S. Bassett, “Tracking the dynamics of the stream of thought reveals its function,” *Technology, Mind, and Behavior*, 2025.
- [376] A. W. Kruglanski and C. Kopetz, *What is so special (and nonspecial) about goals?: A view from the cognitive perspective*. Guilford Press, 2009.
- [377] A. W. Kruglanski, J. Y. Shah, A. Fishbach, R. Friedman, W. Y. Chun, and D. Sleeth-Keppler, “A theory of goal systems,” in *The motivated mind*, pp. 207–250, Routledge, 2018.
- [378] A. L. McGowan, F. Sayed, Z. M. Boyd, M. Jovanova, Y. Kang, M. E. Speer, D. Cosme, P. J. Mucha, K. N. Ochsner, D. S. Bassett, *et al.*, “Dense sampling approaches for psychiatry research: Combining scanners and smartphones,” *Biological Psychiatry*, vol. 93, no. 8, pp. 681–689, 2023.
- [379] K. M. Werner and E. T. Berkman, “Motivational dynamics of self-control,” *Current Opinion in Psychology*, vol. 59, p. 101859, 2024.
- [380] J. B. Rotter, “Generalized expectancies for internal versus external control of reinforcement.,” *Psychological monographs: General and applied*, vol. 80, no. 1, p. 1, 1966.
- [381] E. T. Higgins, “Value from regulatory fit,” *Current directions in psychological science*, vol. 14, no. 4, pp. 209–213, 2005.
- [382] C. S. Dweck and D. S. Yeager, “Mindsets: A view from two eras,” *Perspectives on Psychological science*, vol. 14, no. 3, pp. 481–496, 2019.
- [383] C. S. Dweck, “From needs to goals and representations: Foundations for a unified theory of motivation, personality, and development.,” *Psychological review*, vol. 124, no. 6, p. 689, 2017.
- [384] D. Oyserman, N. A. Lewis Jr, V. X. Yan, O. Fisher, S. C. O’Donnell, and E. Horowitz, “An identity-based motivation framework for self-regulation,” *Psychological Inquiry*, vol. 28, no. 2-3, pp. 139–147, 2017.
- [385] J. M. Zacks, “Event perception and memory,” *Annual review of psychology*, vol. 71, no. 1, pp. 165–191, 2020.

- [386] J. Chen and A. M. Bornstein, “The causal structure and computational value of narratives,” *Trends in Cognitive Sciences*, 2024.
- [387] H. J. Lee, D. M. Smith, C. E. Hauenstein, A. Dworetzky, B. T. Kraus, M. Dorn, D. E. Nee, and C. Gratton, “Precise individual measures of inhibitory control,” *Nature Human Behaviour*, pp. 1–18, 2025.
- [388] H. F. Song, G. R. Yang, and X.-J. Wang, “Training excitatory-inhibitory recurrent neural networks for cognitive tasks: a simple and flexible framework,” *PLoS computational biology*, vol. 12, no. 2, p. e1004792, 2016.
- [389] S. Musall, A. E. Urai, D. Sussillo, and A. K. Churchland, “Harnessing behavioral diversity to understand neural computations for cognition,” *Current opinion in neurobiology*, vol. 58, pp. 229–238, 2019.
- [390] P. I. Jaffe, R. A. Poldrack, R. J. Schafer, and P. G. Bissett, “Modelling human behaviour in cognitive tasks with latent dynamical systems,” *Nature Human Behaviour*, pp. 1–15, 2023.
- [391] J. Achterberg, D. Akarca, D. Strouse, J. Duncan, and D. E. Astle, “Spatially embedded recurrent neural networks reveal widespread links between structural and functional neuroscience findings,” *Nature Machine Intelligence*, vol. 5, no. 12, pp. 1369–1381, 2023.
- [392] L. G. E. Tuleya, *Thesaurus of psychological index terms*. American Psychological Association, 2007.
- [393] G. R. VandenBos, *APA dictionary of psychology*. American Psychological Association, 2007.
- [394] G. R. VandenBos, *APA college dictionary of psychology*. American Psychological Association, 2009.
- [395] D. U. Wulff and R. Mata, “Semantic embeddings reveal and address taxonomic incommensurability in psychological measurement,” *Nature Human Behaviour*, pp. 1–11, 2025.
- [396] J. Friedman, T. Hastie, and R. Tibshirani, “Sparse inverse covariance estimation with the graphical lasso,” *Biostatistics*, vol. 9, no. 3, pp. 432–441, 2008.
- [397] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, *et al.*, “Scikit-learn: Machine learning in python,” *the Journal of machine Learning research*, vol. 12, pp. 2825–2830, 2011.

- [398] M. F. Glasser, T. S. Coalson, E. C. Robinson, C. D. Hacker, J. Harwell, E. Yacoub, K. Ugurbil, J. Andersson, C. F. Beckmann, M. Jenkinson, *et al.*, “A multi-modal parcellation of human cerebral cortex,” *Nature*, vol. 536, no. 7615, pp. 171–178, 2016.
- [399] S. M. Mitchell, S. Lange, and H. Brus, “Gendered citation patterns in international relations journals,” *International Studies Perspectives*, vol. 14, no. 4, pp. 485–492, 2013.
- [400] M. L. Dion, J. L. Sumner, and S. M. Mitchell, “Gendered citation patterns across political science and social science methodology fields,” *Political Analysis*, vol. 26, no. 3, pp. 312–327, 2018.
- [401] N. Caplar, S. Tacchella, and S. Birrer, “Quantitative evaluation of gender bias in astronomical publications from citation counts,” *Nature Astronomy*, vol. 1, no. 6, p. 0141, 2017.
- [402] D. Maliniak, R. Powers, and B. F. Walter, “The gender citation gap in international relations,” *International Organization*, vol. 67, no. 4, pp. 889–922, 2013.
- [403] J. D. Dworkin, K. A. Linn, E. G. Teich, P. Zurn, R. T. Shinohara, and D. S. Bassett, “The extent and drivers of gender imbalance in neuroscience reference lists,” *bioRxiv*, 2020.
- [404] M. A. Bertolero, J. D. Dworkin, S. U. David, C. L. Lloreda, P. Srivastava, J. Stiso, D. Zhou, K. Dzirasa, D. A. Fair, A. N. Kaczkurkin, B. J. Marlin, D. Shohamy, L. Q. Uddin, P. Zurn, and D. S. Bassett, “Racial and ethnic imbalance in neuroscience reference lists and intersections with gender,” *bioRxiv*, 2020.
- [405] X. Wang, J. D. Dworkin, D. Zhou, J. Stiso, E. B. Falk, D. S. Bassett, P. Zurn, and D. M. Lydon-Staley, “Gendered citation practices in the field of communication,” *Annals of the International Communication Association*, 2021.
- [406] P. Chatterjee and R. M. Werner, “Gender disparity in citations in high-impact journal articles,” *JAMA Netw Open*, vol. 4, no. 7, p. e2114509, 2021.
- [407] J. M. Fulvio, I. Akinnola, and B. R. Postle, “Gender (im)balance in citation practices in cognitive neuroscience,” *J Cogn Neurosci*, vol. 33, no. 1, pp. 3–7, 2021.
- [408] D. Zhou, E. J. Cornblath, J. Stiso, E. G. Teich, J. D. Dworkin, A. S. Blevins, and D. S. Bassett, “Gender diversity statement and code notebook v1.0,” Feb. 2020.

- [409] A. Ambekar, C. Ward, J. Mohammed, S. Male, and S. Skiena, “Name-ethnicity classification from open sources,” in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge Discovery and Data Mining*, pp. 49–58, 2009.
- [410] G. Sood and S. Laohaprapanon, “Predicting race and ethnicity from the sequence of characters in a name,” *arXiv preprint arXiv:1805.02109*, 2018.