Exploration versus exploitation in space, mind, and society

Thomas T. Hills[1], **Peter M. Todd**[2], **David Lazer**[3,4,5], **A. David Redish**[6], **Iain D. Couzin**[7,8], and **the Cognitive Search Research Group**[9]

[1]Department of Psychology, University of Warwick, Coventry, UK

[2]Cognitive Science Program, Indiana University, Bloomington, IN, USA

[3]Department of Political Science, Northeastern University, Boston, MA, USA

[4]College of Computer and Information Science, Northeastern University, Boston, MA, USA

[5]Harvard Kennedy School, Harvard University, Cambridge, MA, USA

[6]Department of Neuroscience, University of Minnesota, Minneapolis, MN, USA

[7]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA

[8]Department of Collective Behaviour, Max Planck Institute of Ornithology, Konstanz, Germany

[9]Melissa Bateson (Centre for Behavior and Evolution/Institute of Neuroscience, Newcastle University), Roshan Cools (Radboud University Medical Center, Department of Psychiatry; Donders Institute for Brain, Cognition and Behaviour, Centre for Cognitive Neuroimaging), Reuven Dukas (Department of Psychology, Neuroscience, and Behavior, McMaster University), Luc-Alain Giraldeau (Facultéde Sciences, Université du Qué bec a Montréal), Michael W. Macy (Department of Sociology, Cornell University), Scott E. Page (Center for the Study of Complex Systems, Departments of Political Science and Economics, University of Michigan), Richard M. Shiffrin (Department of Psychological and Brain Sciences, Indiana University), David W. Stephens (Ecology, Evolution, and Behavior, University of Minnesota), and Jeremy W. Wolfe (Brigham and Women's Hospital and Harvard Medical School)

Abstract

Search is a ubiquitous property of life. Although diverse domains have worked on search problems largely in isolation, recent trends across disciplines indicate that the formal properties of these problems share similar structures and, often, similar solutions. Moreover, internal search (e.g., memory search) shows similar characteristics to external search (e.g., spatial foraging), including shared neural mechanisms consistent with a common evolutionary origin across species. Search problems and their solutions also scale from individuals to societies, underlying and constraining problem solving, memory, information search, and scientific and cultural innovation. In summary, search represents a core feature of cognition, with a vast influence on its evolution and processes across contexts and requiring input from multiple domains to understand its implications and scope.

Corresponding author: Hills, T.T. (t.t.hills@warwick.ac.uk).

Exploration versus exploitation

Search, or seeking a goal under uncertainty, is a ubiquitous requirement of life. Animals forage for food, territory, and mates. Humans engage in a wide variety of search behaviors, from looking for lost keys, to finding financial opportunities, to seeking the meaning of existence. Moreover, they search in a wide range of spaces, including visual scenes, memory, social networks, information databases, and more abstract environments such as problem spaces associated with cures for disease, product design, and governmental policy. Search in each of these domains involves trade-offs between exploiting known opportunities and exploring for better opportunities elsewhere. Here we describe how the exploration– exploitation trade-off is fundamental to understanding cognitive behavior at various levels, from its evolutionary origins to the function of cognitive control across domains. How cognitive systems handle this trade-off also has broad implications, from harmful mental disorders to the progress of social innovation. The ubiquity of this trade-off and cognition's solutions to it implicate search as a common framework for understanding many aspects of cognition.

Fields ranging from animal foraging to computer science use different terminology to describe what are fundamentally similar strategies for handling the exploration– exploitation trade-off. In decision-making research, it is referred to as exploitation versus exploration [1,2]; in visual attention, the contrast is on focused versus diffuse search [3,4], in foraging, intensive versus extensive [5], in memory, local versus global [6], and in artificial intelligence, depth-first versus breadth-first search [7].

The diversity of terms used to describe this trade-off reflects a common problem across many cognitive domains (Table 1). In response, numerous solutions have arisen. In computer science, rules that achieve appropriate reinforcement learning have elements of exploitation intermixed with exploration [8]. Ecological search strategies often involve intensive local foraging mixed with occasional exploration phases that move animals from one cluster or region of resources to another [9,10]. Similarly, social insect swarms, such as those of many ants and honeybees, exhibit phases of diffusive and focused search in response to changing resource distributions, as perceived at the level of the colony [11–13].

Traces of the historical significance of this trade-off are found in the evolution of cognition. Recent comparative neurobiological findings of dopaminergic and related mechanisms support a putative common ancestral precursor for foraging behaviors across many animal species. Our developing understanding of the shared control structures among neural correlates of response selection (e.g., pre-frontal cortex and basal ganglia), exploration and memory (e.g., hippocampus), and visual attention (e.g., parietal cortex) has produced the realization that numerous goal-directed processes central to human cognition rely on the integration of search-related architectures. Findings such as these lead to the compelling conclusion that the same cognitive and neural processes underlie much of human behavior involving cognitive search – in both external and internal environments [9,14].

Beyond the individual, our understanding of the exploitation–exploration trade-off extends to the collective behavior of social insects, formal institutions, and populations organized

around political and social identity. In all cases, performance depends on striking an optimal balance between imitating best practices and experimenting with innovations. Here we follow the path from individual to collective behavior, offering a tour of the unifying themes on search that encompass its evolutionary, ecological, neural, algorithmic, and social bases.

From spatial foraging to foraging in mind

Individual organisms must strike the proper balance between global exploration and local exploitation to survive – exploring sufficiently to find resources and exploiting sufficiently to harvest them. This optimal control problem is well-studied in animal foraging, where numerous species focus on finding resources in a cluster (or patch) until their acquisition rate falls below some threshold, as described by the marginal value theorem [15]. At that point, organisms typically leave to seek a new patch [10,16]. Related spatial search strategies apply when there are not clear boundaries around resources. In such cases, predictions can be made about how particular strategies are appropriate to different resource ecologies and cognitive capacities of organisms (Box 1).

Visual search in humans involves at least two nested senses of exploration and exploitation. In a typical visual search task, observers would look for targets among distractor items in a series of scenes or displays. Observers might be said to exploit each scene until the time comes to move on and explore the next scene [17]. On a finer scale, observers explore from item to item in a scene, spending differing amounts of time exploiting each attended item. Figure 1 illustrates some of the factors governing this within-scene exploration. Under reduced-stimulus conditions, observers come close to optimally balancing within-scene exploration and exploitation. Similar issues arise when deciding to move between visual scenes. For example, radiologists may have many images to scan and must balance local search time in each against the need to examine the next radiograph in the stack. The decision to move to the next case is driven by the visual properties of the stimulus [3], the probability of finding a target [18], the reward structure [19], the number of objects to scan in the stimulus [20], and the history of errors [17]. All of these are also factors common to search in other domains.

Modulation between exploration and exploitation is also fundamental to many non-spatial aspects of human behavior. Social search for potential mates (e.g., in marriage and divorce) may proceed from local search (e.g., people highly similar to the searcher) to global search as thresholds for mate quality are adjusted with age [21]. Humans searching for information in the external environment – information foraging (e.g., on the Web or through social networks) – also switch adaptively between local and global search: people leave a local patch of web pages when they perceive that its value has fallen below what can be found globally elsewhere [22]. In decision making, the search for cues on which to base choices often starts with those that have been most useful in the past (local focus) and proceeds to others until enough cues have been found to select an option [23]. This pattern also arises in problem solving. For example, when engaged in an anagram task that involves making words from a set of random letters, people use the discovery of past solutions to determine how long to stay in a local 'patch' [24], as they also do when angling for fish in a sequence of virtual ponds on a computer screen [25]. Expecting resources near where (or when) other

resources have been found may be a broadly adaptive cognitive bias – explaining, for example, the tendency for people and nonhuman primates to expect 'streaky' outcomes, sometimes called the hot-hand fallacy [26].

Exploration can be random, treating all courses of action more or less equally, or guided by beliefs about the structure of the environment [27–29]. The former, often called model-free, is driven by stimulus-response relationships and defines exploration as an increased probability of choosing an outcome associated with a lower expected value or of choosing any behavior with unknown consequences. Belief- guided search, often called model-based, involves a cognitive representation or map of the relational structure of the environment. Model-based decisions can inform both where and when to search by guiding search to regions that provide more information. An example of model-based 'when' search is exploring to see whether a cake is finished baking only after sufficient time has passed to allow it to do so. In the model-free case, there is no model of the temporal dynamics of baking, simply knowledge that baking cakes need checking. While model-free exploration is typically defined by choosing outcomes with lower expected rewards than exploitation, model-based exploration can lead to greater long-term rewards by using cognitive representations to decide both where and when exploration is most likely to pay dividends in the future [30,31]; from this perspective, learning can be conceived of as a foraging process [32].

Are internal and external search processes governed by common underlying mechanisms? Figure 2 illustrates a recent priming study that found that experience with more local exploitation in a spatial foraging task led to more focused production of word solutions in a subsequent anagram task [33]. The common mechanism is also supported by findings demonstrating that measures of executive capacity (e.g., working memory span) are related to search in various domains, both internal and external [34–37], including visual search [38,39]. Furthermore, dynamic switching between local and global search has been suggested to account for human decisions to move from one 'region' of memory to another – for example, when recalling types of animal in a verbal fluency task [6] – as well as decisions to leave visible patches (e.g., berry bushes in a simulated berry-picking task) [40].

The neurobiology of cognitive search

Cognitive search involves the capacity to maintain goals in stable neural representations and to relax these representations as needed to create opportunities for flexibility and exploration in either the internal or external world. These capacities allow us to search not only in the present, but also in the past and the future (Box 2). Research in neuroscience is uncovering core neural architectures for guiding search and mediating the exploration– exploitation trade-off – stability and flexibility – in relation to goal maintenance.

A host of neural circuits initiate and control search based on goals and motivations and the simultaneous identification of uncertainty with respect to their attainment. with the exploration–exploitation trade-off and its requisite goal maintenance and updating are the dorsolateral and medial prefrontal cortices [41,42]. These regions interact with numerous other brain structures associated with the control of attention, action evaluation, and out-

come prediction, including the anterior cingulate cortex, the hippocampal formation, and the dorsal, ventral, medial and lateral aspects of the striatum. For example, the prefrontal areas interact with the hippocampal formation more predictably during flexible behaviors in rats [43–45] and humans [46,47] and this is reduced as habits form and dorsal and lateral aspects of the striatum learn action chains [48–50].

Although these interactions are governed by numerous neuromodulators, a key component is dopamine [51,52]. Dopamine is known to facilitate attentional control in response to unexpected stimuli [53], to modulate reward- seeking across species [9,54], and to manipulate goal maintenance and updating via phasic and tonic activation patterns [55,56]. How does dopamine mediate stability and flexibility in the brain? Dopamine receptor stimulation is proposed to increase the signal-to-noise ratio of neuronal firing [57]. This leads to increased robustness of goal representations in the face of intervening distractors [58], maintaining the stability of goals and actions associated with their exploitation.

In addition to dopamine, the stabilization of search with respect to goal-relevant representations may also involve acetylcholine and noradrenaline (norepinephrine), which modulate levels of activation associated with expected and unexpected uncertainty [59], respectively – that is, whether the uncertainty originates from cues that are known to be unreliable or cues that are grossly inconsistent with the current context. In particular, noradrenaline has been proposed to mediate the shift from exploration to exploitation by changing the tonic background activity of neurons in the prefrontal cortex [60,61], thereby altering the threshold for new goal representations to compete for activation. When one course of action leads to less-than-expected rewards, neuromodulators in the brain lower the threshold for exploring alternatives [40,44]. Similarly, serotonin may also contribute to modulating search and, in particular, the propensity for model-based versus model-free search strategies [28,62].

The role of these brain regions and neurotransmitters naturally leads to predictions about pathologies of goal- directed and, specifically, search-related behavior associated with corresponding lesions or dysfunctions of neuromodulation. Many symptoms of Parkinson's disease, addiction, obsessive–compulsive disorder, schizophrenia, attention deficit disorder, and depression can be interpreted as deficits in search-related neural architectures and fall at the extremes of the exploration–exploitation trade-off [9,62,63].

Collective problem solving in social species

Social systems also face the exploration–exploitation trade-off, but they mediate this trade-off in two distinct ways. Some social organisms, such as honeybees, and many ant species, delegate exploration and exploitation across different individuals so that individuals do not have to navigate the trade-off. Beginning with Karl von Frisch [64], biologists recognized that social insect colonies are often divided into scouts who explore independently for new food sources, and workers who are recruited to exploit what others find (e.g., by pheromones in ants or the waggle dance in honeybees). The exploratory behavior of scouts is also modulated by dopaminergic agonists, but it is unclear to what extent this is homologous with the mechanisms described above [65]. Individuals also exhibit the

flexibility to switch roles, with those failing to find recruitment signals within the colony switching to scouting behavior [66]. Consequently, the proportion of a colony participating in exploration increases if there are not enough existing foraging locations for the workers to exploit. Across species, the capacity for individuals to exploit resources discovered by others creates an additional exploration– exploitation trade-off beyond that posed for individuals – one that is a strength and a weakness of social systems (Box 3).

A second way that social systems can negotiate this trade-off is by altering the underlying connectivity between individuals and thus manipulating the ability of individuals to socially influence one another. The optimal level of connectivity also depends uniquely on the structure of the search problem. For search over smooth environments with single peaks, experiments on humans and simulations have found that fully connected groups can outperform less well-connected groups; however, as environments became more multipeaked and rugged, less well-connected communication networks lead to longer periods of exploration and better long-term solutions [67,68]. Such diffuse communication systems may also resemble the foraging tactics of groups who can track moving resource gradients that are detectable collectively but are undetectable to individual foragers (Figure 3).

Both social–cognitive approaches to the exploration– exploitation trade-off – mixed roles across individuals and controlled communication connectivity between individuals – are found in cultural systems of human innovation. Federal systems rely on separate states to explore new policies and exploit what other states find [69]. Franchises experiment with new products and advertising plans in distinct locations and then carefully control the diffusion of these innovations [70]. Also, firms directly control roles and imitation by attempting to balance the exploration of new product-market domains with an exploitative focus on increasing existing product-market efficiency; firms better able to achieve this see higher sales growth [71]. Roles and connectivity can also be extended over time, to adapt to changing environments. For example, search for innovations by groups (including industries and nations) can be improved through balancing small bursts of social learning during periods of environmental change against long stretches of exploiting known strategies during periods of environmental stability [72]. In environments where solutions are hard to find but nonetheless indicated by other solutions, imitation can act as a form of group memory and enhance long-term performance by facilitating search around good solutions that have been found in the past [73]. This is not unlike the search behavior observed among schooling fish [74].

In self-organized social domains, sometimes called coordination problems, people can self-assign roles to achieve the level of diversity in exploration and exploitation needed to track past success while simultaneously seeking better solutions [75]. Beneficial diversity can also come from differences in individual ability, but its usefulness depends on how information is aggregated across individuals and the kinds of problems they are solving [13,76,77]. Furthermore, exploration can be facilitated by dissent and ignorance: dissent can delay collective decisions, increasing the number of options considered by keeping exploratory discussion open [78], while individuals who participate in collective decisions with unbiased or uninformed preferences may tend to inhibit strongly opinionated minorities and promote

a fair representation of opinions within groups [79]. Finally, in philosophy and political science, group deliberation is often viewed as a process of social search through a problem space in which different institutional rules affect the balance between exploring that space and exploiting the existing positions of individuals [80]. Thus, the capacity for different individuals to have different amounts and qualities of information extends the adaptive achievement of exploration–exploitation trade-offs from social foragers to more cognitively abstract domains of culture and policy.

Concluding remarks and future directions

This review provides a framework for envisioning cognition as a search process characterized by the exploration– exploitation trade-off, which scales from individuals to society (Figure 4). Viewing cognition as a search process highlights the potential conserved properties and problems, shared neural and cognitive mechanisms, and homologous or convergent evolutionary origins of many components of cognition [81]. Moreover, some components of cognitive search appear to be domain general, possibly facilitating the goal maintenance necessary for model-based internal foraging over varieties of cognitive repre-sentations [33].

Socially interacting groups regularly engage in search, acting as distributed cognitive systems to the extent that they collectively develop a cognitive representation of the environment that no one individual fully contains and behave based on that representation [82]. Consequently, they clearly face the same exploration–exploitation trade-off, which they can solve in novel ways such as altering the exploratory biases of individuals or altering the communication structure between individuals. Cognitive systems at the individual level utilize similar mechanisms. For example, individuals can alternate between exploration and exploitation in space and brains can alternate between exploration and exploitation by modulating connectivity via changes in tonic firing. Considering links between the individual and group levels invites questions about the extent to which social adaptations for the exploration– exploitation trade-off are built on cognitive faculties for mediating the trade-off within individuals. If spatial foraging provides a precursor for internal foraging, it may also have provided pre-adaptations for aspects of socially coordinated exploration.

The cognitive search perspective also requires us to consider how cognitive representations of different types of search space – whether internal or external – interact with different possible search mechanisms (Box 4). Do some cognitive representations obey the same rules as physical 2D or 3D space or can those representations increase the dimensionality of search problems, in much the same way as genomic complexity increases the dimensionality of evolutionary search [83]? Does this increased dimensionality require new search strategies? A more complete understanding of cognitive search requires studies of the mechanisms for controlling and optimizing search as well as of how they interact with the structure of various internal and external environments. This will include understanding how search processes respond to both dynamic and noisy environments [84,85]. We also need to develop methodologies for disentangling changes in cognitive representations from changes in the search processes that navigate them. For example, is age-related cognitive decline in

memory a consequence of problems with the search processes [37] or with changes in the memory representation, which may reflect degradation [86] or enrichment [87], or both?

Search also offers a framework for understanding cognitive problems in relation to their molecular, genetic, and neural underpinnings. Knowing how these building blocks influence search at one level of description can offer insights into why we see changes at other levels. As noted above, maladaptive states of both individual and group search lie at the extremes of too much exploitation (compulsiveness, perseveration, and groupthink) or too much exploration (impulsiveness, inattentiveness, and failure to leverage social information). Understanding how these extremes arise in response to age, pathology, genetics, environment, and social structure represents significant challenges to a wide range of fields in cognitive science – challenges for which a shared conceptual framework offers both insights and a basis for communicating our ideas and questions.

References

1. Cohen J, et al. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. Philos. Trans. R. Soc. Lond. B: Biol. Sci. 2007; 362:933. [PubMed: 17395573]

2. March JG. Exploration and exploitation in organizational learning. Organ. Sci. 1991; 2:71–87.

3. Treisman AM, Gelade G. A feature-integration theory of attention. Cogn. Psychol. 1980; 12:97–136. [PubMed: 7351125]

4. Wolfe JM, et al. Guided search: an alternative to the feature integration model for visual search. J. Exp. Psychol. Hum. Percept. Perform. 1989; 15:419–433. [PubMed: 2527952]

5. Benhamou S. How many animals really do the Lévy walk? Ecology. 2007; 88:1962–1969. [PubMed: 17824427]

6. Hills TT, et al. Optimal foraging in semantic memory. Psychol. Rev. 2012; 119:431–440. [PubMed: 22329683]

7. Korf RE. Depth-first iterative-deepening: an optimal admissible tree search. Artif. Intell. 1985; 27:97–109.

8. Sutton, RS.; Barto, AG. Reinforcement Learning: An Introduction. MIT Press; 1998.

9. Hills TT. Animal foraging and the evolution of goal-directed cognition. Cogn. Sci. 2006; 30:3–41. [PubMed: 21702807]

10. Stephens, DW., et al. Foraging: Behavior and Ecology. University of Chicago Press; 2007.

11. Bonabeau, E., et al. Swarm Intelligence: From Natural to Artificial Systems. Oxford University Press; 1999.

12. Deneubourg JL, et al. Random behaviour, amplification processes and number of participants: how they contribute to the foraging properties of ants. Physica D. 1986; 22:176–186.

13. King AJ, et al. Is the true ''wisdom of the crowd'' to copy successful individuals? Trends Cogn. Sci. 2009; 13:197–200.

14. Hills, TT.; Dukas, R. The evolution of cognitive search. In: Todd, PM., et al., editors. Cognitive Search: Evolution, Algorithms, and the Brain. MIT Press; 2012. p. 11-24.

15. Charnov EL. Optimal foraging, the marginal value theorem. Theor. Popul. Biol. 1976; 9:129–136. [PubMed: 1273796]

16. Stephens, DW.; Krebs, JR. Foraging Theory. Princeton University Press; 1987.

17. Chun MM, Wolfe JM. Just say no: how are visual searches terminated when there is no target present? Cogn. Psychol. 1996; 30:39–78. [PubMed: 8635311]

18. Wolfe JM, et al. Rare items often missed in visual searches. Nature. 2005; 435:439–440. [PubMed: 15917795]

19. Kristjansson A, et al. Fortune and reversals of fortune in visual search: reward contingencies for pop-out targets affect search efficiency and target repetition effects. Atten. Percept. Psychophys. 2010; 72:1229–1236. [PubMed: 20601703]

20. Cousineau D, Shiffrin RM. Termination of a visual search with large display size effects. Spat. Vis. 2004; 17:327–352. [PubMed: 15559108]

21. Hills T, Todd PM. Population heterogeneity and individual differences in an assortative agent-based marriage and divorce model (MADAM) using search with relaxing expectations. J. Artif. Soc. Social Simul. 2008; 11:5.

22. Pirolli P. Rational analyses of information foraging on the web. Cogn. Sci. 2005; 29:343–373. [PubMed: 21702778]

23. Gigerenzer, G., et al. Efficient cognition through limited search. In: Todd, PM.; Gigerenzer, G., editors. Ecological Rationality: Intelligence in the World. Oxford University Press; 2012. p. 608

24. Wilke A, et al. Fishing for the right words: decision rules for human foraging behavior in internal search tasks. Cogn. Sci. 2009; 33:497–529. [PubMed: 21585478]

25. Hutchinson JMC, et al. Patch leaving in humans: can a generalist adapt its rules to dispersal of items across patches? Anim. Behav. 2008; 75:1331–1349.

26. Scheibehenne B, et al. Expectations of clumpy resources influence predictions of sequential events. Evol. Hum. Behav. 2011; 32:326–333.

27. Dayan P, Berridge KC. Model-based and model-free Pavlovian reward learning: revaluation, revision, and revelation. Cogn. Affect. Behav. Neurosci. 2014; 14:473–492. [PubMed: 24647659]

28. Blanco NJ, et al. The influence of depression symptoms on exploratory decision-making. Cognition. 2013; 129:563–568. [PubMed: 24055832]

29. Otto AR, et al. Physiological and behavioral signatures of reflective exploratory choice. Cogn. Affect. Behav. Neurosci. 2014; 14:1167–1183. [PubMed: 24664860]

30. Markant DB, Gureckis TM. Is it better to select or to receive? Learning via active and passive hypothesis testing. J. Exp. Psychol. Gen. 2014; 143:94–122. [PubMed: 23527948]

31. Sims CR, et al. Melioration as rational choice: sequential decision making in uncertain environments. Psychol. Rev. 2013; 120:139. [PubMed: 23230889]

32. Metcalfe J, Jacobs J. People's study time allocation and its relation to animal foraging. Behav. Processes. 2010; 83:213–221. [PubMed: 20026197]

33. Hills TT, et al. The central executive as a search process: priming exploration and exploitation across domains. J. Exp. Psychol. Gen. 2010; 139:590–609. [PubMed: 21038983]

34. Hills TT, Pachur T. Dynamic search and working memory in social recall. J. Exp. Psychol. Learn. Mem. Cogn. 2012; 38:218–228. [PubMed: 21859235]

35. Rakow T, et al. Biased samples not mode of presentation: re-examining the apparent underweighting of rare events in experience-based choice. Organ. Behav. Hum. Decis. Process. 2008; 106:168–179.

36. Unsworth N. Individual differences in working memory capacity and episodic retrieval: examining the dynamics of delayed and continuous distractor free recall. J. Exp. Psychol. Learn. Mem. Cogn. 2007; 33:1020–1034. [PubMed: 17983310]

37. Hills T, et al. Mechanisms of age-related decline in memory search across the adult life span. Dev. Psychol. 2013; 49:2396–2404. [PubMed: 23586941]

38. Sobel K, et al. Individual differences in working memory capacity and visual search: the roles of top-down and bottom-up processing. Psychon. Bull. Rev. 2007; 14:840–845. [PubMed: 18087947]

39. Anderson DE, et al. A common discrete resource for visual working memory and visual search. Psychol. Sci. 2013; 24:929–938. [PubMed: 23572280]

40. Wolfe JM, Danielson J. Visual foraging behavior: when are the berries riper on the other side of the screen? J. Vis. 2012; 12:265.

41. Fuster, J. The Prefrontal Cortex. Elsevier; 2008.

42. Goldman-Rakic PS. Cellular basis of working memory. Neuron. 1995; 14:477–485. [PubMed: 7695894]

43. Jones MW, Wilson MA. Theta rhythms coordinate hippocampal–prefrontal interactions in a spatial memory task. PLoS Biol. 2005; 3:e402. [PubMed: 16279838]

44. Hyman JM, et al. Working memory performance correlates with prefrontal–hippocampal theta interactions but not with prefrontal neuron firing rates. Front. Integr. Neurosci. 2010; 4:1–13. [PubMed: 20161992]

45. Benchenane K, et al. Coherent theta oscillations and reorganization of spike timing in the hippocampal–prefrontal network upon learning. Neuron. 2010; 66:921–936. [PubMed: 20620877]

46. Voss JL, et al. Hippocampal brain-network coordination during volitional exploratory behavior enhances learning. Nat. Neurosci. 2010; 14:115–120. [PubMed: 21102449]

47. Schacter DL, Addis DR. On the nature of medial temporal lobe contributions to the constructive simulation of future events. Philos. Trans. R. Soc. Lond. B: Biol. Sci. 2009; 364:1245–1253. [PubMed: 19528005]

48. Barnes TD, et al. Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. Nature. 2005; 437:1158–1161. [PubMed: 16237445]

49. Smith KS, Graybiel AM. A dual operator view of habitual behavior reflecting cortical and striatal dynamics. Neuron. 2013; 79:361–374. [PubMed: 23810540]

50. Dezfouli A, Balleine BW. Habits, action sequences and reinforcement learning. Eur. J. Neurosci. 2012; 35:1036–1051. [PubMed: 22487034]

51. Robbins T, Roberts A. Differential regulation of fronto-executive function by the monoamines and acetylcholine. Cereb. Cortex. 2007; 17:i151–i160. [PubMed: 17725997]

52. Cools R, D'Esposito M. Inverted-U-shaped dopamine actions on human working memory and cognitive control. Biol. Psychiatry. 2011; 69:e113–e125. [PubMed: 21531388]

53. Schultz W. Predictive reward signal of dopamine neurons. J. Neurophysiol. 1998; 80:1–27. [PubMed: 9658025]

54. Barron AB, et al. The roles of dopamine and related compounds in reward-seeking behavior across animal phyla. Front. Behav. Neurosci. 2010; 4:1–9. [PubMed: 20126432]

55. Redish AD, et al. Reconciling reinforcement learning models with behavioral extinction and renewal: implications for addiction, relapse, and problem gambling. Psychol. Rev. 2007; 114:784–805. [PubMed: 17638506]

56. Beeler JA, et al. Tonic dopamine modulates exploitation of reward learning. Front. Behav. Neurosci. 2010; 4:1–14. [PubMed: 20126432]

57. Servan-Schreiber D, et al. A network model of catecholamine effects: gain, signal-to-noise ratio, and behavior. Science. 1990; 249:892–895. [PubMed: 2392679]

58. Durstewitz D, Seamans JK. The dual-state theory of prefrontal cortex dopamine function with relevance to catechol-*O*-methyltransferase genotypes and schizophrenia. Biol. Psychiatry. 2008; 64:739–749. [PubMed: 18620336]

59. Yu AJ, Dayan P. Uncertainty, neuromodulation, and attention. Neuron. 2005; 46:681–692. [PubMed: 15944135]

60. Aston-Jones G, Cohen JD. An integrative theory of locus coeruleus–norepinephrine function: adaptive gain and optimal performance. Annu. Rev. Neurosci. 2005; 28:403–450. [PubMed: 16022602]

61. Kehagia AA, et al. Learning and cognitive flexibility: frontostriatal function and monoaminergic modulation. Curr. Opin. Neurobiol. 2010; 20:199–204. [PubMed: 20167474]

62. Huys QJM, et al. Bonsai trees in your head: how the Pavlovian system sculpts goal-directed choices by pruning decision trees. PLoS Comput. Biol. 2012; 8:e1002410. [PubMed: 22412360]

63. Winstanley, CA., et al. Search, goals, and the brain. In: Todd, PM., et al., editors. Cognitive Search: Evolution, Algorithms, and the Brain. MIT Press; 2012. p. 125-156.

64. Frisch von, K. The Dance Language and Orientation of Bees. Harvard University Press; 1967.

65. Liang ZS, et al. Molecular determinants of scouting behavior in honey bees. Science. 2012; 335:1225–1228. [PubMed: 22403390]

66. Beekman M, et al. What makes a honeybee scout? Behav. Ecol. Sociobiol. 2007; 61:985–995.

67. Mason WA, et al. Propagation of innovations in networked groups. J. Exp. Psychol. Gen. 2008; 137:422–433. [PubMed: 18729708]

68. Lazer D, Friedman A. The network structure of exploration and exploitation. Admin. Sci. Quart. 2007; 52:667–694.

69. Kollman K. Decentralization and the search for policy solutions. J. Law Econ. Organ. 2000; 16:102–128.

70. Sorenson O, Sorensen JB. Finding the right mix: franchising, organizational learning, and chain performance. Strat. Mgmt. J. 2001; 22:713–724.

71. He Z-L, Wong P-K. Exploration vs. exploitation: an empirical test of the ambidexterity hypothesis. Organ. Sci. 2004; 15:481–494.

72. Rendell L, et al. Why copy others? Insights from the social learning strategies tournament. Science. 2010; 328:208–213. [PubMed: 20378813]

73. Wisdom TN, et al. Social learning strategies in networked groups. Cogn. Sci. 2013; 37:1383–1425. [PubMed: 23845020]

74. Berdahl A, et al. Emergent sensing of complex environments by mobile animal groups. Science. 2013; 339:574–576. [PubMed: 23372013]

75. Roberts ME, Goldstone RL. Adaptive group coordination and role differentiation. PLoS ONE. 2011; 6:e22377. [PubMed: 21811595]

76. Page, SE. The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies. Princeton University Press; 2007.

77. Krause S, et al. Swarm intelligence in humans: diversity can trump ability. Anim. Behav. 2011; 81:941–948.

78. Nemeth CJ, Wachtler J. Creative problem solving as a result of majority vs minority influence. Eur. J. Soc. Psychol. 1983; 13:45–55.

79. Couzin ID, et al. Uninformed individuals promote democratic consensus in animal groups. Science. 2011; 334:1578–1580. [PubMed: 22174256]

80. Habermas, J.; McCarthy, T. A Theory of Communicative Action. Beacon Press; 1985.

81. Todd, PM., et al., editors. Cognitive Search: Evolution, Algorithms, and the Brain (Strüngmann Forum Reports). MIT Press; 2012.

82. Goldstone RL, Gureckis TM. Collective behavior. Top. Cogn. Sci. 2009; 1:412–438. [PubMed: 25164995]

83. Gavrilets S. Evolution and speciation on holey adaptive landscapes. Trends Ecol. Evol. 1997; 12:307–312. [PubMed: 21238086]

84. Gureckis TM, Love BC. Learning in noise: dynamic decision-making in a variable environment. J. Math. Psychol. 2009; 53:180–193. [PubMed: 20161328]

85. Otto A, et al. Navigating through abstract decision spaces: evaluating the role of state generalization in a dynamic decision-making task. Psychon. Bull. Rev. 2009; 16:957–963. [PubMed: 19815805]

86. Borge-Holthoefer J, et al. Modeling abnormal priming in Alzheimer's patients with a free association network. PLoS ONE. 2011; 6:e22651. [PubMed: 21829639]

87. Ramscar M, et al. The myth of cognitive decline: non-linear dynamics of lifelong learning. Top. Cogn. Sci. 2014; 6:5–42. [PubMed: 24421073]

88. Nonacs P. State dependent behavior and the marginal value theorem. Behav. Ecol. 2001; 12:71–83.

89. Beachly WM, et al. On the economics of sit-and-wait foraging: site selection and assessment. Behav. Ecol. 1995; 6:258–268.

90. Kareiva P, Odell G. Swarms of predators exhibit 'preytaxis' if individual predators use area-restricted search. Am. Nat. 1987; 130:233–270.

91. Viswanathan G, et al. Optimizing the success of random searches. Nature. 1999; 401:911–914. [PubMed: 10553906]

92. Hills TT, et al. Adaptive Lévy processes and area-restricted search in human foraging. PLoS ONE. 2013; 8:e60488. [PubMed: 23577118]

93. Ferreira AS, et al. The influence of the environment on Lévy random search efficiency: fractality and memory effects. Physica A. 2012; 391:3234–3246.

94. Klein R. Inhibitory tagging system facilitates visual search. Nature. 1988; 334:430–431. [PubMed: 3405288]

95. Peterson MS, et al. Visual search has memory. Psychol. Sci. 2001; 12:287–292. [PubMed: 11476094]

96. Wolfe JM, et al. Attention is fast but volition is slow. Nature. 2000; 406:691. [PubMed: 10963584]

97. Buckner RL, Carroll DC. Self-projection and the brain. Trends Cogn. Sci. 2007; 11:49–57. [PubMed: 17188554]

98. Redish, AD. The Mind within the Brain: How We Make Decisions and How those Decisions Go Wrong. Oxford University Press; 2013.

99. Kurth-Nelson Z, et al. A theoretical account of cognitive effects in delay discounting. Eur. J. Neurosci. 2012; 35:1052–1064. [PubMed: 22487035]

100. Otto AR, et al. Working-memory capacity protects model-based learning from stress. Proc. Natl. Acad. Sci. U.S.A. 2013; 110:20941–20946. [PubMed: 24324166]

101. Genovesio A, et al. Prefrontal–parietal function: from foraging to foresight. Trends Cogn. Sci. 2014; 18:72–81. [PubMed: 24378542]

102. Tolman EC. The determiners of behavior at a choice point. Psychol. Rev. 1938; 45:1–41.

103. Johnson A, Redish AD. Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. J. Neurosci. 2007; 27:12176–12189. [PubMed: 17989284]

104. Schacter DL, Addis DR. The cognitive neuroscience of constructive memory: remembering the past and imagining the future. Philos. Trans. R. Soc. Lond. B: Biol. Sci. 2007; 362:773–786. [PubMed: 17395575]

105. Hassabis D, et al. Patients with hippocampal amnesia cannot imagine new experiences. Proc. Natl. Acad. Sci. U.S.A. 2007; 104:1726–1731. [PubMed: 17229836]

106. Keramati M, et al. Speed/accuracy trade-off between the habitual and the goal-directed processes. PLoS Comput. Biol. 2011; 7:e1002055. [PubMed: 21637741]

107. Johnson A, et al. The hippocampus and exploration: dynamically evolving behavior and neural representations. Front. Hum. Neurosci. 2012; 6:216. [PubMed: 22848196]

108. van der Meer M, et al. Information processing in decision-making systems. Neuroscientist. 2012; 18:342–359. [PubMed: 22492194]

109. Daw ND, et al. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat. Neurosci. 2005; 8:1704–1711. [PubMed: 16286932]

110. van Bergen Y, et al. Nine-spined sticklebacks exploit the most reliable source when public and private information conflict. Proc. Biol. Sci. 2004; 271:957–962. [PubMed: 15255051]

111. Barnard CJ, Sibly RM. Producers and scroungers: a general model and its application to captive flocks of house sparrows. Anim. Behav. 1981; 29:543–550.

112. Vickery WL, et al. Producers, scroungers, and group foraging. Am. Nat. 1991; 137:847–863.

113. Giraldeau, L-A.; Caraco, T. Social Foraging Theory. Princeton University Press; 2000.

114. Giraldeau LA, et al. Potential disadvantages of using socially acquired information. Philos. Trans. R. Soc. Lond. B: Biol. Sci. 2002; 357:1559–1566. [PubMed: 12495513]

115. Simons AM. Many wrongs: the advantage of group navigation. Trends Ecol. Evol. 2004; 19:453–455. [PubMed: 16701304]

116. Guttal V, Couzin ID. Social interactions, information use, the evolution of collective migration. Proc. Natl. Acad. Sci. U.S.A. 2010; 107:16172–16177. [PubMed: 20713700]

117. Dussutour A, et al. The role of multiple pheromones in food recruitment by ants. J. Exp. Biol. 2009; 212:2337–2348. [PubMed: 19617426]

118. Deneubourg JL, et al. Probabilistic behaviour in ants: a strategy of errors? J. Theor. Biol. 1983; 105:259–271.

119. Weidenmüller A, Seeley TD. Imprecision in waggle dances of the honeybee (Apis mellifera) for nearby food sources: error or adaptation? Behav. Ecol. Sociobiol. 1999; 46:190–199.

120. Galton F. Vox populi. Nature. 1907; 75:450–451.

Box 1. Animal foraging

How long should a foraging animal stay within a given resource patch? In Charnov's (1976) marginal value theorem, a forager with perfect knowledge of the environment spends t time units exploiting and depleting a resource patch and takes t time units to travel from one patch to another. Within each patch, the forager gains an amount of food specified by the function $g(t)$, with the rate of gain decelerating the more the patch has been exploited. The mean intake rate for an animal exploiting a sequence of these patches is:

$$\frac{g(t)}{(\tau+t)} \qquad \text{[I]}$$

The value of *t* with the highest possible intake rate satisfies the 'marginal value' condition:

$$g'(t)=\frac{g(t)}{(\tau+t)} \qquad \text{[II]}$$

This means that the forager should leave a patch when its instantaneous rate of intake in the patch $g'(t)$ falls to the mean intake rate for the environment as a whole. Hence, foragers should leave patches relatively early in rich environments but leave later in poorer environments. Qualitatively, this prediction is well supported [10,88].

Strategies change with increasing uncertainty about the quality of the current resource or the likelihood of finding something better. With non-depleting patches, a forager should 'stay forever' in a patch of the best type once it finds it; for example, trap-building spiders occupy permanent feeding sites after a period of sampling. This is a type of optimal stopping problem, comparing the value of the current resource to the forager's expectations about future rewards (cf. [15] versus [89]).

In situations with poorly defined patch boundaries (e.g., a bumblebee exploiting clumps of flowers), foragers often rely on area-restricted search, intensifying their search in a local area after encounters with food but exploring more as food encounters decrease [9,90]. This can be modeled as satiety S controlling frequency of turning r with an increasing function $r = R(S)$. Now, turning will be highest in areas of highest prey density. Such a process can generate 'preytaxis', in which foragers climb gradients of prey abundance [90].

At larger spatial scales, resource gradients can be more complex; gradient-following rules may then fail by becoming trapped on inferior local maxima far from global peaks of resource abundance. Making occasional large movements independent of recent experience may then be beneficial. One strategy that accomplishes this is a Lévy flight, which is operationally defined as any heavy-tailed search strategy such that path lengths, *l*, are drawn from a power-law distribution, with $P(I_j)=I_j^{-\mu}$ and $1 < \mu < 3$ [91]. Under

what conditions Lévy flights are common or optimal, however, is hotly debated [5,92,93].

Box 2. From foraging to deliberation

Cognitive search in humans offers the capacity to search within one's knowledge of the external world via simulation of that external world. This creates the capacity for deliberation. When faced with a complex and difficult choice, humans will deliberate over their options by constructing imagined future outcomes and evaluating them serially [97,98]. This is a search through an internal model of the world [99–101].

In the 1930s, Tolman observed that rats would sometimes pause and reorient back and forth at choice points and suggested that this behavior entailed a form of 'vicarious trial and error' – an internal search process [102]. It is now known that, during these behaviors, hippocampal representations encode future options [103] and hippocampal–prefrontal interactions increase [45,104], a process called episodic future thinking. Humans with damage to the hippocampal–prefrontal interacting system do not deliberate over choices and do not show these internal search processes [46,105].

The same trade-offs that impact external search processes affect this type of internal search. For example, should one continue to take time to explore possibilities (exploration) or should one stop and take the best choice found so far (exploitation)? Because the internal search is a computation, it takes time, which will slow one's interaction with the external world. Interestingly, there is strong evidence that mammalian decision making includes internal (cognitive) search, external (interactive) search processes, and cached (exploitation) processes [63]. Which process drives decision making depends on the time available for introspection and the information known about the world [106–108]. The advantage of internal search is that possible outcomes can be considered without the time commitment and inherent danger of implementing them. However, taking too much time to deliberate can be disastrous. At some point, the search (exploration) has to be stopped and the action (exploitation) taken. When one is truly time limited (for example, in sports), this internal, deliberative search process is often purposely bypassed by a less flexible (cached) action-selection system that has been learned through extensive training, often referred to as overlearned or automatic processing [108,109].

Box 3. Social search

Many organisms utilize social information when searching for resources (such as food or an appropriate habitat) and when making decisions. In this way individuals can integrate personally-acquired information about their environment with social information, the latter based on the positions and behavior of others [74,110]. If effort is required to find items, organisms may limit their costs by exploiting the search effort of others. Thus, individuals can 'scrounge' information from 'producers' [111]. If too many individuals adopt a scrounging strategy, however, it pays to invest in personal search. This leads to a frequency-dependent balance of strategies within populations, although this often does not maximize population-level resource acquisition due to the game-theoretic nature of producer– scrounger behavior [111–114].

For search scenarios, organisms must detect – and climb – noisy, long-range environmental (e.g., temperature, salinity, resource) gradients. Here, social interactions can provide substantial additional benefit by allowing individuals, simply through grouping, to average their imperfect estimates of temporal and spatial cues (the so-called 'wisdom-of-crowds' effect) [115]. Due to the investment necessary to obtain personal information, however, this again sets the scene for producers (searchers) to be exploited by others [116]. Alternatively, a recent experimental study of gradient detection in schooling fish discovered that distributed sensing, correlated with group size, facilitates adaptive responses to complex environmental gradients [74]. Thus, groups can exhibit a size-dependent exploratory awareness of the environment that is not present at the individual level.

In other organisms, such as the eusocial insects (e.g., ants, bees), additional communication strategies enhance search. Many ants deposit chemical trails during search, allowing others to determine which areas have previously been searched and where food has been found, along with (approximately) when, as indicated by the degree of trail evaporation. Thus, different chemicals employed simultaneously can guide potential searchers to engage in rapid, directed motion through areas of low interest and intensive search of other areas [117]. Similarly, honeybees modulate the search efforts of other group members through the waggle dance, which, performed within the nest, indicates to others the approximate direction, quality, and distance to resources. As with other search processes, some degree of error has been found to be adaptive in the dance, allowing an effective balance between exploration and exploitation [118,119].

Box 4. Outstanding questions

- What cognitive representations of search spaces are there, what are their features, how do we switch between them, and how might they change over time?

- How do attentional processes guide information search in external and internal environments and how might this be related to other cognitive domains, such as categorization, learning, or language production?

- How are search processes biologically and cognitively adapted to the structure of different environments?

- Does role switching in social systems operate in the same way as role-switching decisions in the context of individual foraging? In addition, does a similar kind of group-level working memory capacity govern the exploration–exploitation trade-off in social systems?

- Can differences in exploration policies between social insect groups, franchises, and federal systems be linked to aspects of their internal structure?
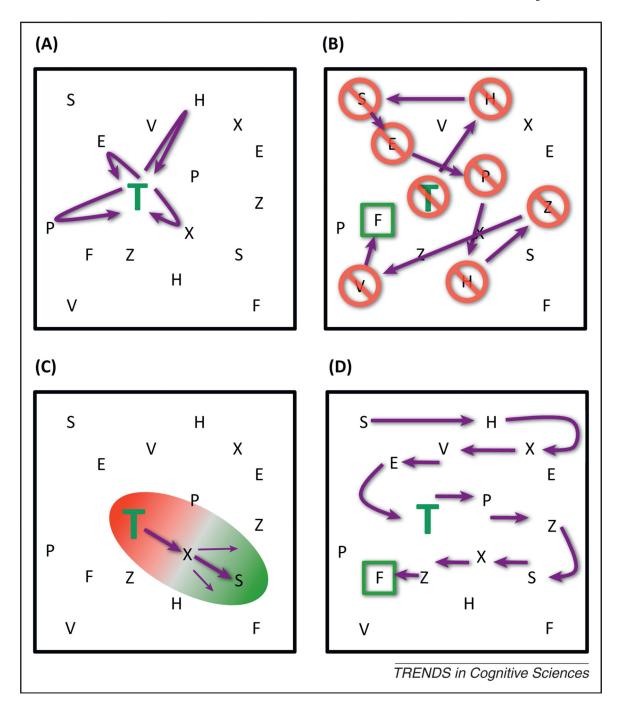
Figure 1.

Local exploration in human visual search. In visual search, people look for specific letter targets; for example, Fs in an array of letter distractors. The local exploitation step is the act of recognizing a letter and determining whether it is your target. The local exploration step is the act of selecting the next letter (accomplished at a rate of about 20–40 letters/s). All else being equal, visual attention is drawn to salient items in the field (**A**). How then are we to avoid perseverating on one incorrect but vivid letter? One answer (**B**) is to rely on the phenomenon of 'inhibition of return' (IOR) [94]. If one attends or fixates on an item and

then deploys one's gaze or attention away from that item, it becomes harder to bring the gaze or attention back to the original item than to move it elsewhere. It was originally believed that IOR would permit attention to sample the display without replacement. Unfortunately, further research found that visual search was not markedly impaired when IOR was blocked and observers had to sample with replacement. A more moderate view might hold that IOR serves to bias exploration toward new items (**C**) even if it does not absolutely prevent return to a rejected item. However, given enough time, observers can adopt strategies that allow them to prevent perseveration [95]. Thus, for example, one might 'read' a display from side to side and top to bottom. This more controlled, prospective strategy (**D**) would avoid sampling with replacement but would slow the rate with which items can be processed. In some cases, a more chaotic strategy will get you to the target more quickly [96].
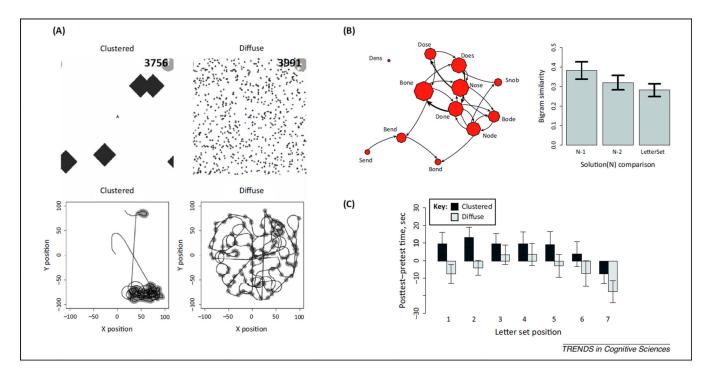
Figure 2.

Priming from external to internal search. (**A**) People initially searched for hidden targets (invisible to the participants, but shown in black) in environments with either clustered or diffuse resource distributions. The lower two panels show typical foraging patterns in the two environments. Following this period of foraging in external space, people's internal search behavior was quantified in a lexical search task. In the lexical search task, participants were asked to find 30 words across a series of letter sets, with multiple words possible per letter set. (**B**) A network representation of search in one letter set where participants found multiple words in the letter set NSBDOE. Nodes represent solutions and links between nodes represent transitions between solutions. Node size is proportional to the number of participants who provided that solution. Link thickness is proportional to the number of participants who made that transition. Bigram similarity between previous solutions and the original letter set shows a clear tendency to produce solutions that are more like the previous solution, N − 1, than like solutions two items back, N − 2, or the letter set. Priming is shown in panel (**C**), with individuals in the clustered condition staying longer in letter sets (post-test–pretest time) than individuals in the diffuse condition. Reproduced from [33].
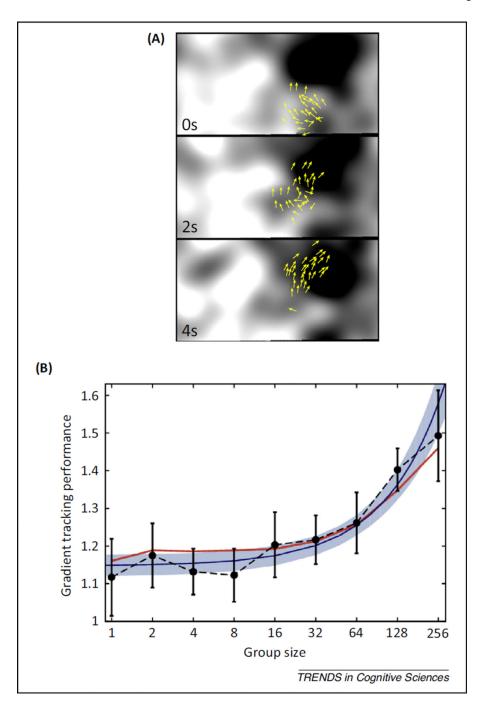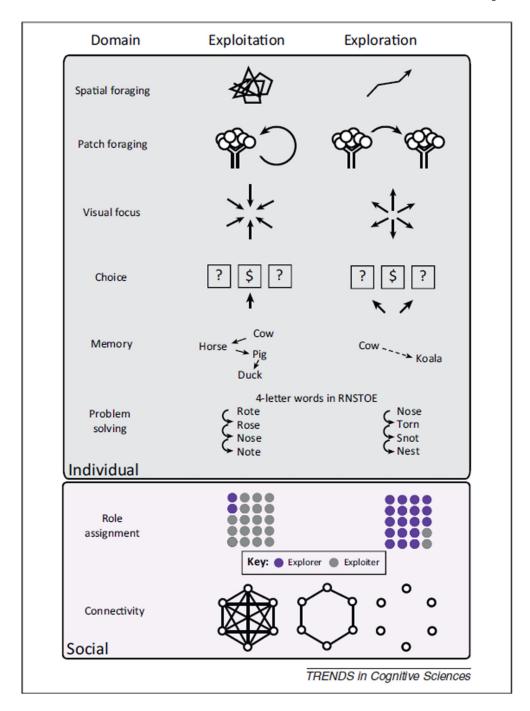
Figure 3.
Collective search allows individuals to track environmental gradients that occur at long length scales relative to the size of individuals, as well as to track gradients that they cannot individually detect. (**A**) The positions of 30 fish are shown as they negotiate, collectively, a moving light field toward the preferred darker regions. The snapshots are 2 s apart, with time progressing downward. (**B**) Performance, as measured by the time-averaged darkness level at fish locations, increases with group size. Experiments (points connected by dotted line) and simulations (red line) show this to be a function of individual fish simply

manipulating their speed (moving slower in darker regions) while maintaining connectivity via locally mediated social interactions. Thus, counter to the conventional view of the 'wisdom of crowds', in which individuals pool imperfect estimates [120], here detection of the local light gradient is absent at the individual level (individuals employ only a scalar measure of the light intensity at their present location), yet exploration and exploitation emerge as a dynamic property of the collective. Reproduced from [74].

Figure 4.
Examples of cognitive approaches to the exploration–exploitation trade-off, as discussed in the text.

Table 1

Examples of trade-offs between exploitation and exploration across cognitive domains

Animal foraging	Exploiting a known berry bush versus exploring for new bushes
Visual search	Analyzing one spot on a chest radiograph versus looking for the next spot to check
Information search	Searching within a document versus searching for new documents; deciding when to accept an item on a menu versus continuing to look for new items
Search in memory	Trying to remember more African animals versus switching to Australian animals
Search in problem solving	Focusing on solutions that have worked in the past versus seeking new solutions
Social (group) learning	Learning or copying existing knowledge versus using innovation to seek new knowledge