

The best of both worlds: Dual systems of reasoning in animals and AI

Mark Kelly^a, Andrew B. Barron^{b,*}

^a Opteran Technologies, Sheffield Innovation Centre, 217 Portobello, Sheffield S1 4DP, United Kingdom

^b Macquarie University, Balaclava Rd, Macquarie Park, NSW 2109, Australia

ARTICLE INFO

Keywords:

Selective attention
Degeneracy
Mushroom body
Modularity
Neural network
Monte-Carlo tree search

ABSTRACT

Much of human cognition involves two different types of reasoning that operate together. Type 1 reasoning systems are intuitive and fast, whereas Type 2 reasoning systems are reflective and slow. Why has our cognition evolved with these features? Both systems are coherent and in most ecological circumstances either alone is capable of coming up with the right answer most of the time. Neural tissue is costly, and thus far evolutionary models have struggled to identify a benefit of operating two systems of reasoning. To explore this issue we take a broad comparative perspective. We discuss how dual processes of cognition have enabled the emergence of selective attention in insects, transforming the learning capacities of these animals. Modern AIs using dual systems of learning are able to learn how their vast world works and how best to interact with it, allowing them to exceed human levels of performance in strategy games. We propose that the core benefits of dual processes of reasoning are to narrow down a problem space in order to focus cognitive resources most effectively.

1. Introduction

Dual process theories of reasoning posit that much of our cognition is comprised of two distinct processes or forms of reasoning; Type 1 processes are fast, automatic, and intuitive while Type 2 processes are slower, more deliberative, and reasoned (Osman, 2004; Slovic, 1996; Kahneman, 2011; Wason & Evans, 1974; Evans & Stanovich, 2013). Here we briefly summarise the various dual process theories, and consider how and why dual processes of reasoning have evolved. We explore the cognitive advantages conferred by dual-process systems, particularly in relation to attention and learning.

Thus far humans have dominated discussion of dual systems of reasoning, with some Type 2 systems considered as unique to human cognition. Here we take a very broad comparative perspective to consider the evolution of this cognitive feature. We argue that dual processes of reasoning are not unique to humans, nor is the general approach of multi-process control unique to reasoning. We draw from new studies of animal cognition that suggest dual systems of reasoning are widespread and have a long evolutionary history. We also draw from new developments in artificial intelligence (AI) which have converged on dual systems of reasoning as an effective solution to complex problems.

Type 1 processes are associative or intuitive and are clearly very widespread in animals (Perry, Barron, & Cheng, 2013). While the Type 2

processes that demand human-level abstract linguistic abilities are likely unique to humans in the animal kingdom, not all forms of Type 2 processes require language. Behaviours reminiscent of rule-based reasoning, meta-cognition, forward planning, and abstract reasoning are found within a wide range of animals, including insects (Avergues-Weber & Giurfa, 2013; Cope et al., 2018; Giurfa, Zhang, Jenett, Menzel, & Srinivasan, 2001; Menzel & Fischer, 2011; Perry & Barron, 2013; Perry & Chittka, 2019; ten Cate & Healy, 2017). The mechanisms needed to support some Type 2 processes such as working memory, attention, representation, and internal predictive models are also present in many animals (de Bivort & van Swinderen, 2016; Perry, Barron, & Chittka, 2017; Nityananda, 2016; Cross, Carvell, Jackson, & Grace, 2020; Kamdayi & Osvath, 2017; ten Cate & Healy, 2017; Menzel & Fischer, 2011; Solvi, Al-Khudhairy, & Chittka, 2020).

Insects have proven useful model systems for dissecting the mechanisms of various forms of cognition (Perry et al., 2013, 2017), and here we discuss how simpler forms of dual processes of cognition occur in their miniature and more tractable brains. We discuss what features of Type 1 and Type 2 processes are present in insects, and how they might have evolved.

Some artificial intelligences have converged on motifs present in dual process systems of cognition. A particularly telling example comes from the family of AIs developed by Google Deepmind to challenge superhuman levels of performance in strategy games like go and chess

* Corresponding author at: Macquarie University, 14 Eastern Road, Science Road, Macquarie Park, NSW 2109, Australia.

E-mail address: andrew.barron@mq.edu.au (A.B. Barron).

<https://doi.org/10.1016/j.cognition.2022.105118>

Received 10 February 2021; Received in revised form 29 March 2022; Accepted 1 April 2022

Available online 19 April 2022

0010-0277/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

(Hassabis, Kumaran, Summerfield, & Botvinick, 2018; Silver et al., 2016, 2017; Silver et al., 2018). These are comprised of both associative deep neural networks and forward modelling tree-search systems, with interactions between them. The interacting dual systems support a form of attention in such a way that their combined performance is greater than either system could achieve alone (Silver et al., 2016). This dual-systems approach to AI has delivered new levels of performance, and new cognitive capacities.

The analogy to dual processes of reasoning is compelling. Despite AIs being designed rather than evolved systems they can still help us understand the potential benefits of dual systems of reasoning for performance, computational speed, and efficiency. These are powerful evolutionary constraints in most biological systems as nervous tissue and activity is so metabolically costly (Sterling & Laughlin, 2015).

How a dual process cognitive system could have evolved has not received a lot of attention. The comparative perspective we present here sheds new light on the issue by considering this architecture of reasoning as forms of degeneracy and multi-process control; frameworks which emphasises robustness, flexibility, and evolvability (Maleszka, Mason, & Barron, 2014; Whitacre & Bender, 2010).

2. Dual processes of reasoning in humans

Models of human cognition that posit two systems of reasoning have been discussed since the early 20th Century (Osman, 2004). The different theories generally agree that Type 1 processes are fast and intuitive or automatic, while Type 2 processes are slower, reflective, and involve more effortful deliberation (Osman, 2004; Barbey & Sloman, 2007; Sloman, 1996; Evans & Stanovich, 2013; Evans, 2019).

Theories primarily differ in how the systems interact; the levels of influence each has over the other and how conflict between them is mediated (Osman, 2004; Evans, 2019). Different dual process theories have been reviewed by Osman (2004) and a summary is given in Table 1.

Many theories postulate roles for consciousness in the two systems. Some view Type 1 process as sub- or pre-conscious and Type 2 as conscious, while others argue consciousness exists across the two processes as and when required (Osman, 2004). Dual process theories of cognition are not predicated on consciousness. For this article we focus on dual process systems of cognition that are either non-conscious (AI) or possibly minimally conscious (insects). This allows a focus on the potential benefits of dual process systems irrespective of the question of consciousness.

Evans & Stanovich argue that Type 1 and Type 2 processes of reasoning are best thought of as networks of systems, rather than as each being a single structure (Evans & Stanovich, 2013). Evans builds a picture of combinations of cognitive resources able to be assembled for different specific Type 1 or Type 2 systems as required (Evans, 2019). In this way dual process theories of reasoning are reconcilable with some of the modular theories of mind (Erana, 2012; Coltheart, 1999; Evans, 2019; Carruthers, 2006). It is perhaps more tractable to see how additional cognitive modules could evolve within an already present dual

systems framework, than to imagine a whole new form of reasoning evolving de-novo (Erana, 2012).

Theories developed by Evans & Over postulate that Type 1 processes act automatically to highlight representations of relevant parts of the world using simple heuristics. Representations which progress into Type 2 processes then undergo more analytical analyses (Evans & Over, 2013; Osman, 2004). Here we see a more sequential element to interactions between the two systems, with Type 1 operating first, and subsequently directing Type 2 responses.

Others like Sloman (1996) and Stanovich and West (2000) argue that the two systems function more in parallel, with each contributing distinctly different ways of processing information about the world. Operating in parallel does not however preclude the two processes interacting. Parallel operation requires mechanisms to resolve conflicts between the two types of systems when they cue contradictory responses (Denes-Raj & Epstein, 1994).

Viewing Type 1 and Type 2 processes as collections of cognitive systems allows us to entertain multiple different ways for the systems to interact, both in parallel and series. Default interventionist models propose that rapid Type 1 processes generate a 'default' response, which may or may not be modified by intervening slower Type 2 processes (Evans & Stanovich, 2013; De Neys & Glumicic, 2008). Often in these models Type 2 processes are cast as inhibitory of action, putting the brakes on an inappropriate response generated by Type 1 processes (Gronchi & Giovannelli, 2018). Hybrid two-stage models propose a continually active monitoring process able to detect conflicts between the outcomes of Type 1 and Type 2 processes, and shift cognitive resources to Type 2 processes if a conflict is found. (De Neys & Glumicic, 2008; Thompson, 2013; Newell, Lagnado, & Shanks, 2015).

Evans and Over (2013) and Stanovich and West (2000) argue that Type 2 processes are influential in learning some Type 1 responses. Here, automatic skills and reasoning abilities are first learnt explicitly by Type 2 processes before being transferred to Type 1 through persistent practice (Osman, 2004). It is likely that all of these different methods of interaction could occur in different networks across the brain.

Interactions between two systems and their consequences have been modelled and explored theoretically. Trimmer et al. (2008) built a dual system with a fast process reacting instantaneously to the current input and a slower process accumulating data to increase confidence. Models differed in whether or not the two processes ran in parallel and, if so, what interactions were present. Both systems could perform well separately, but optimal performance was achieved when the two systems ran in parallel but were aware of the other's response (Trimmer et al., 2008).

Stanovich (2012) expanded upon dual processing theory by introducing a tripartite model in which Type 2 processes involved two separate systems; an algorithmic system for modelling the world and simulating outcomes, and a reflective system for goal prioritisation and beliefs. These are then coupled to the Type 1 processes which govern autonomous behaviours. Interactions occur between the three systems. Type 1 processes act as a pre-attentive stage to Type 2. The reflective system of Type 2 can then initiate simulations within the algorithmic mind and, if deemed necessary, override the outcomes of Type 1

Table 1
Comparison of features ascribed to Type 1 and Type 2 processes and how they interact for different dual process theories.

	Type 1 features	Type 2 features	Interaction
Evans & Over	Preconscious and based on heuristics, associative, selects potentially relevant stimuli	Analytic, explicit, sequential, conscious and reflective	Type 2 operates on representations highlighted by Type 1
Sloman	Intuitive, limited to experienced phenomenon	Conscious rule interpretation, can combine representations in novel ways	Parallel but interaction, Type 2 can override Type 1
Stanovich & West	Automatic and autonomous, subconscious, context dependent, based on heuristics	Controlled and analytical, decoupled from context	Functionally distinct and parallel systems, Type 2 endorses or overrides Type 1
Tversky & Kahnemann	Intuitive, informal, unstructured, based on heuristics	Controlled and deliberate, rule-based, flexible	Type 2 processes act as a regulator of Type 1 responses
Epstein	Emotionally driven, encodes experiences	Non-experiential and rational, abstract, analytical	Functionally distinct and parallel systems, Type 2 processes can endorse or override Type 1

systems. (Stanovich, 2012). It is important to note here that while three systems are present, they are still divided into two processes, with Type 2 processes involving two systems, and Type 1 involving a single system.

Allen Newell presented his vision for a Unified Theory of Cognition, the Soar architecture, which splits processing into three systems (Newell, 1990). The most basic system is parallel, bottom-up, and automatic. A second system then acts as a bottleneck, selecting a single action from those generated automatically. These two systems approximately correspond to Type 1 processes, while a third system acts when knowledge is incomplete and inferences must be made (Newell, 1990). Again we can see the importance of attention in how information from one process interacts with the other, here acting as it's own system.

This distillation into just two processes has been a common criticism of dual process theory, with arguments made that this binary dichotomy does not adequately explain the range of processes identified in human reasoning (Osman, 2004). Proposals for four systems – implicit heuristic, implicit rule-based, explicit heuristic, and explicit rule-based – have been made (Moshman, 2000), as have models in which reasoning is a single system operating on a continuum between intuitive and analytical extremes (Hammond, 1996; Osman, 2004).

Milli, Lieder, and Griffiths (2017) sought to shed light upon why we might have only two processes by considering cognitive systems from the perspective of a bounded optimality problem. In a simple two-choice experiment, increasing the variance of the inputs saw computational models comprised of multiple processes outperform simpler models. Additional processes come with additional costs however, both in building and maintaining the processes themselves and the infrastructure required to unify them into a single output (Sterling & Laughlin, 2015). These costs place an upper bound on the number of processes (Milli et al., 2017).

Comparing speed-accuracy trade-off, Milli et al. (2017) argued that two processes provided optimal performance across a wide range of plausible input characteristics. In these optimised two process models, one was found to be fast and error-prone while the other was slow and accurate. This is not entirely unexpected, as when two processes run it is inevitable that one will be faster than the other and, given the constraint of the speed-accuracy trade off, the faster system will likely be more error-prone.

What is common to all dual process theories, however, is that the distinction between the processes is more significant than one fast and erroneous and one slow. The two processes are different kinds of reasoning. To explore why cognition should have both a simple fast form of reasoning and a more complex slower form of reasoning we consider examples of dual process systems that have developed outside of human cognition. We can accept that dual systems of reasoning are highly diverse, but why should cognition involve two processes at all, let alone two processes that run at different speeds?

For humans several benefits have been proposed for dual-process systems of reasoning that range from post-hoc justification of a choice to oversight and corrective action for a wrong choice (Evans, 2019). Any of these might apply in theory, but while it is possible to construct problems that can only be solved with type 2 reasoning, for many ecologically relevant situations Type 1 reasoning intuitively the correct answer faster (Evans, 2019). In these situations the role of Type 2 processing seems to be to rationalise or justify a choice already made.

If one process is sufficient to reach a decision in most ecological circumstances, why evolve and again engage two? Tomlin (Tomlin, Rand, Ludvig, & Cohen, 2015) found no intrinsic benefit to a slow deliberate cognitive process over a fast automatic process in a series of simple evolutionary models. Any advantage of a deliberative process was undermined by its slower speed and greater costs. Neural tissue is hard and costly to build and metabolically expensive to operate (Sterling & Laughlin, 2015). Evolution is littered with examples of animals evolving to minimise neural tissue whenever possible (Sterling & Laughlin, 2015). Maintaining and operating two systems for decision making seems like an indulgence that is hard to explain.

In summary, human dual process theories embrace a diversity of cognitive systems within each type, a diversity of modes of interaction, and a diversity of functions. Currently theories are rather quiet on the evolutionary benefits of applying to a problem two distinct processes of reasoning that run at different speeds. For a different perspective on this we discuss advances in AI which have converged on dual processes of reasoning to outperform humans in strategy games.

3. Dual processes of reasoning in artificial intelligence

AlphaGo famously beat humans in the strategy game of Go (Silver et al., 2016), and we argue that its success came from use of dual process cognitive systems (Schrittwieser et al., 2020; Silver et al., 2017; Silver et al., 2018). AlphaGo and its many improvements differ in details and implementation, but all utilise a fast deep neural network paired with a slower tree search approach (Schrittwieser et al., 2020; Silver et al., 2016, 2017; Silver et al., 2018).

A tree search works by taking the current state of the game board and modelling the outcomes of all possible moves until the end of the game and then selecting moves most likely to result in a win. For the game of Go however, the number of possible moves is astronomically large and it is too computationally intensive to simply brute-force the solution. Monte-Carlo methods reduce this search by only exploring a subset of random possible moves which are constrained to give a good representation of the input space.

In parallel to the tree search, a deep neural network was trained based on feedback from the moves selected and played. This network was unable to simulate future positions in the game, but was able to assign a policy value to particular board configurations (Silver et al., 2016). While the deep neural network learned the correlation between reward and the game's spatial layout, the graph network was able to explore the game in the temporal domain and, in doing so, established a more causal relationship between a move and a particular game state.

Here we see a clear analogy to the intuitive Type 1 vs. analytical Type 2 processes of reasoning. Deep neural networks have long been criticised for being black boxes lacking any true comprehension or reasoning for their decisions. They function associatively by distilling large volumes of training data into a near instantaneous decision, similar to intuition and Type 1 processes. The graph network and tree search act similar to Type 2 processes by modelling the world as a series of consequences of actions. They are also more comprehensible, since they identify why a move is beneficial by analysis of the subsequent moves and states it can lead to. This is similar to how human Go masters employ a look-ahead strategy to mentally model the consequences of possible moves. It is rule based, reflective, requires simulation of possible outcomes, and is relatively slow and computationally intensive (Silver et al., 2016).

Whilst AlphaGo's systems could function separately, optimal performance that exceeded human levels of play was achieved only when they were allowed to interact (Silver et al., 2016). AlphaGo utilised its deep neural network's ability to rapidly assess board configurations and constrain the Monte-Carlo tree search onto states more likely to be beneficial, drastically increasing the expected return of the search. This mode of operating shares similarities with proposals by Stanovich (2011) and Evans and Over (2013) that Type 1 processes function as pre-attentive stages, forming rapid associations to select potentially relevant representations for further Type 2 processing.

Moves determined beneficial by the tree search can then be played and the feedback used to train the fast neural network (Silver et al., 2016) and over time the neural network begins to improve its mimicry of the tree search. This harkens back to ideas around implicit learning where decisions or actions first calculated by Type 2 processes can be transferred into Type 1 through practice and repetition (Osman, 2004). With the wider gamut of moves selected by the tree search, more informed play is observed producing stable learning when trained purely through self-play (Silver et al., 2018).

To enable this tree search, AlphaGo was pre-programmed with the

rules of the game at hand. A later variant, MuZero, instead employed an additional structure within its Type 2 processing dedicated to building models of the world for simulating future states (Schrittwieser et al., 2020). This is close to the motif proposed in Stanovich's tripartite model in which one kind of Type 2 processing is an algorithmic capacity dedicated to simulating the likely outcomes of possible decisions (Stanovich, 2011).

The similarities between AlphaGo and dual processing has been discussed previously by Anthony, Tian, and Barber (2017). Here they present a similar model with its slower processes acting to plan out new actions, and its deep neural network updating to gradually generalise across those plans. The fast system was again used to guide slower tree searches and increase efficiency, which in turns provides better training data for the deep neural network (Anthony et al., 2017). Once again we can see this almost symbiotic relationship between the two systems based on attention and learning.

These examples from modern AI show that a major benefit of dual processes of reasoning has been the emergence of the capacity to learn complex tasks and rules governing how the world works. Dual processes of reasoning have given AI the ability of learning to learn by allowing their Type-2 processes to create and explore internal models of the external environment. Type 1 processes provide a form of an attentional system that narrows down the problem space and allows efficient exploration of internal models. This is a transformative step-change in cognitive capacity and we would argue it to be one of the fundamental benefits of a dual systems approach.

This provides a suggestion *why* dual processes of reasoning have evolved in human cognition. But AIs are designed and built systems. Analogy with AI will not help us understand how dual processes of reasoning might evolve. To understand *how* they evolved, we need to consider more biological examples.

4. Dual process systems in insects

Obviously we are not suggesting that insects reason like humans, but their cognition is complex enough for them to be interesting. Most insects studied have very robust associative learning, and some insects (particularly ants, bees, and wasps) are capable of various forms of non-elemental learning (Perry et al., 2017). They can learn basic abstract relations such as same and different (Giurfa et al., 2001), and rule-based relationships like transitive inference (Tibbetts, Agudelo, Pandit, & Riojas, 2019). They are capable of basic forms of forward modelling (Perry & Chitka, 2019), abstract representations of stimuli (Solvi et al., 2020), and attentional processing (de Bivort & van Swinderen, 2016; Nityananda, 2016; Paulk et al., 2014; Wiedenman & O'Carroll, 2013).

Experiments have suggested that bumblebees initially solve problems conceptually before utilising a heuristics-based approach after repeated trials (MaBouDi, Solvi, & Chitka, 2020). Bumblebees were trained using sucrose reward to prefer stimuli that were either larger or smaller than their comparator. Over the first 20 training trials bees showed evidence of learning this relational rule, but in later training most bees switched and adopted a win-stay / lose-switch strategy to the patterns (MaBouDi et al., 2020). This still afforded the animals an effective solution to the task without requiring complex comparisons, allowing for a glimpse into how Type 1 and Type 2 processes could approach the same problem. In transfer tests where the heuristic failed, the bees reverted back to solving via comparing sizes (MaBouDi et al., 2020).

Insect brains are remarkably modular, with processing divided into distinct functional regions with a large diversity in connections and connection types (Shih et al., 2015). Here we will focus on the mushroom bodies. The mushroom bodies are large with nearly a third of the bee's million neurons residing within them (Strausfeld, Hansen, Li, Gomez, & Ito, 1998). Two exist within each brain, one on either side, receiving multi-modal sensory input and transforming that into trainable valence values (Aso et al., 2014; Strube-Bloss, Nawrot, & Menzel,

2011). The dimensionality of the input is then drastically increased with dozens of inputs connecting to thousands of its constituent neurons (Aso, Sitaraman, et al., 2014). Each of these neurons respond selectively to combinations of inputs, typically possessing a low background firing rate (Strube-Bloss et al., 2011). In addition to their roles in associative learning, the mushroom bodies have been shown to play roles in navigation (Kamhi, Barron, & Narendra, 2020), selective attention (de Bivort & van Swinderen, 2016; Nityananda, 2016; van Swinderen et al., 2009), context generalisation (Liu, Wolf, Ernst, & Heisenberg, 1999), sleep control (Sitaraman, Aso, Rubin, & Nitabach, 2015), habituation (Brembs, 2009), and regulating fat storage (Al-Anzi & Zinn, 2018). Here we will focus on their roles in learning and selective attention.

Despite being typically seen as centres of olfactory learning (Aso, Sitaraman, et al., 2014), flies with temporarily silenced mushroom bodies can still form and recover context dependent long term olfactory memories (Zhao et al., 2019). Flies are capable of learning to avoid a specific odour without the mushroom bodies, but only displayed this behaviour if the training and testing arenas were visually identical; the odour memory was dependent on the background visual context. Here we think that they are displaying simple snapshot memory; upon punishment or reward a snapshot of all sensory input is captured, associating all contiguous inputs with a given contingency. With functioning mushroom bodies, flies attended selectively to the tight temporal correlation between odour and shock. This yielded learning of the odour only and not the background.

We argue that the mushroom bodies did not provide an entirely new form of learning to the insect brain, but instead added a form of attention to an existing associative learning framework. Other studies have highlighted how the mushroom bodies allow insects to learn about individual stimuli within the environment and also the relationships between stimuli (Cope et al., 2018; Devaud, Papouin, Carcaud, Sandoz, & Giurfa, M., 2015). Flies are able to distinguish between distinct objects in their environment, and to judge if an upstream odour mixture emanates from one source or two (Sehdev, Mohammed, Triphan, & Szyszka, 2019). Visual neuron responses also suggest an ability to distinguish foreground objects from the background (Aptekar, Keleş, Lu, Zolotova, & Frye, 2015).

The mushroom bodies are the more recently evolved structure of the arthropod brain (Strausfeld et al., 1998). It appears that the mushroom body dependent pathway has added a capacity for selective attention to insect cognition, and with that came the ability to learn specific features allowing learning of individual features independent of background.

Studies of the mushroom bodies of the cockroach have found evidence for two parallel but interacting streams within them (Takahashi, Nishino, Domae, & Mizunami, 2019). The two streams are controlled by different types of feedback neurons, and behaved independently but had a level of interaction. Neurophysiological recordings in honey bees have shown that mushroom body output occurs in two phases, with an initial response observed to all stimuli and a later response that is selectively activated for odours recently associated with reward (Strube-Bloss et al., 2011). This initial activity was observed to be incredibly fast; so fast it's outputs reached maximal activity and began to subside before the outputs of the antennal lobe (an input structure to the mushroom bodies) had done likewise (Strube-Bloss, Herrera-Valdez, & Smith, 2012).

It is suggested that this signal from the mushroom bodies provides a fast modulatory response to the other parts of the brain, potentially steering processing onto select representations (Strube-Bloss et al., 2012). Neural recordings in flies have shown how output from the mushroom body are capable of affecting the sensitivity of sensory processing neurons (Hu, Zhang, & Wang, 2010) and could allow for quick impressions to be fed back into the sensory pathway for isolating out objects of interest (Hu et al., 2010; Strube-Bloss et al., 2012). It is not currently known if the earlier phase of activity has any influence over the latter, but given their close proximity we think it likely some interaction would occur. Studies in cockroaches found some level of interaction between the parallel systems they observed, but no

indication of its function (Takahashi et al., 2019).

Thanks to the fruit fly's extensive genetic toolkit, the three lobes of their mushroom bodies has been mapped in exquisite detail. One lobe, the α'/β' or prime lobe responds rapidly, while the others, the α/β and γ , or non-prime lobes, are slower to respond. The prime lobe is also the most excitable and fastest to respond, having an uncharacteristically high baseline activity (Groschner, Chan Wah Hak, Bogacz, DasGupta, & Miesenböck, 2018; Inada, Tsuchimoto, & Kazama, 2017). Parts of the non-prime lobes have been observed to be much more selective in their responses, integrating information more precisely over longer periods of time (Groschner et al., 2018).

We will consider the prime and non-prime lobes in flies as separate channels and similar to the two phase response seen in bees (Strube-Bloss et al., 2011) and the two streams observed in cockroaches (Takahashi et al., 2019). We argue that the prime lobes act in a manner analogous to Type 1 processes and the non-prime lobes are analogous to a Type 2 process. While separate, there are numerous connections between the lobes. In particular, gap junctions connect the prime lobe up to each of the non-prime lobes (Hervé & Derangeon, 2013) adding a capability for near instantaneous transference of information between the lobes. Two large interneurons with roles in memory consolidation and sparsity also span across the entire mushroom body (Lin, Bygrave, De Calignon, Lee, & Miesenböck, 2014; Amin, Apostolopoulou, Suárez-Grimalt, Vrontou, & Lin, 2020; Wu, Shih, Lee, & Chiang, 2013; Liu et al., 2016). Recurrent circuits utilising neurons outside the mushroom body also exist, allowing activity in one lobe to affect the others (Aso et al., 2014).

Instinctive startle responses to either a novel odour (Hattori et al., 2017) or tapping on the container (Sun et al., 2017) required processing within the prime lobes. Co-activation of the α/β non-prime lobes acted to suppress this startle response. This appears similar to dual process theories of Slovic and Epstein in which Type 2 processes are seen as inhibitory of Type 1, and able to override their innate responses (see Table 1). The γ lobes did not behave analogously however, acting to endorse the startle response rather than override it (Sun et al., 2017)

whereas dual process theories would argue it should also suppress (Stanovich, 2012).

Learning appears to be particularly sensitive to γ lobe activity (Krashes, Keene, Leung, Armstrong, & Waddell, 2007), with responses from the γ lobe driving the association of stimuli with both food (Al-Anzi & Zinn, 2018; Tsao, Chen, Lin, Yang, & Lin, 2018) and water (Lin, Oswald, et al., 2014). The γ lobes were not required for the innate following of a trail of water vapour however. This innate behaviour relied on the prime lobes (Lin, Oswald, et al., 2014). Similarly, while not needed to associate stimuli with food, the prime lobes controlled the innate approach and consumption of food (Al-Anzi & Zinn, 2018). Here we can see how basic, instinctive behaviours like following water vapour trails or approaching food are controlled by the prime lobes, while non-prime lobes are needed for learning about them.

Some output neurons of the mushroom body feed directly back into other compartments, creating a multi-layered output. As can be seen in Fig. 1, these all occur within the slower system. The prime lobe operates more as a simple feedforward network, similar to what was seen in the deep neural networks of AlphaGo (Silver et al., 2016). One of these feedback neurons within the non-prime lobes has been found to play a role in integrating satiety signals and preventing approach of food related stimuli when the fly is already well fed (Krashes et al., 2009; Tsao et al., 2018). This is analogous to some dual process theories positing that Type 2 processes integrate high level goal prioritisation and act to override inappropriate Type 1 responses.

A multi-layered output like that of the non-prime lobes would take time to establish, slowly building as information propagates through the layers. This acts more sequentially, and requires selective attention to ensure activity is focused on stable representations throughout the activity's establishment. Conversely, the prime lobes are capable of responding near instantaneously to their inputs, allowing for rapid if limited analysis.

The mushroom bodies of insects appears to operate as a dual system, with some key similarities and differences to dual processes of reasoning. Activity within the fast prime lobes governs instinctive

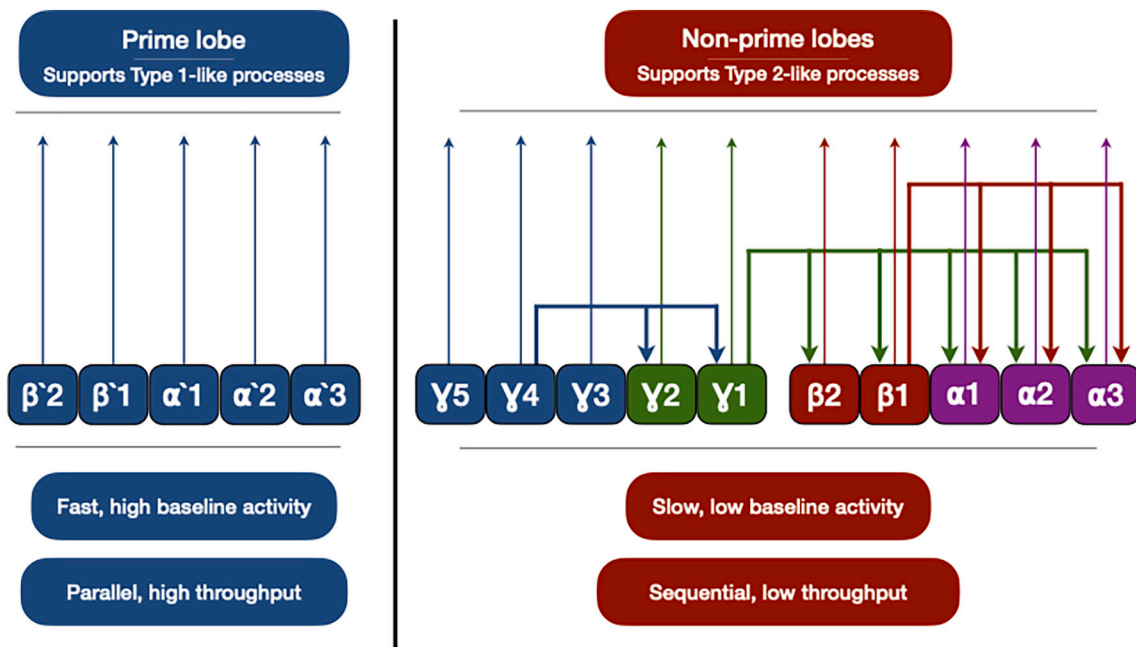


Fig. 1. Diagram (adapted from Fig. 17c of Aso, Hattori, et al. (2014)) showing output neuron connections and characteristics for the three lobes of a fruit fly's mushroom body, separated into two channels. The first channel, supporting Type-1 processes, is comprised of the α'/β' lobe, or prime lobe, and has no direct feedback connections. This channel is significantly more active at rest than the others, and its neuronal responses are rapid and brief (Inada et al., 2017). The other channel, comprised of the α/β and γ lobes, or non-prime lobes, feeds its outputs back to create a multilayered output system and is typically quieter and slower to respond (Inada et al., 2017). This multi-layered readout takes time to establish as each feedback signal is integrated, but is able to incorporate things like internal state in its calculations. (Krashes et al., 2009).

behaviours such as startle responses and consumption, while activity in the slower non-prime lobes can override this behaviour based on factors like internal state or aid them through learning.

We also see processing in the mushroom body supporting a form of selective attention such that flies are able to learn associations between specific stimuli and responses so that background stimuli are ignored. The insects provide another example of dual-process systems narrowing down a problem space to refine what is learned and how it is learned such that cognitive performance is improved. The selective attention capacity supported by the mushroom body is a very basic form of learning-to-learn as it filters what matters for learning from background. The precise associations learned by the mushroom body are an elementary form of modelling the external world, such that learned relationships recognised in the external world can override innate behavioural responses. Whether or not the insect brain uses purely correlations to form associations or can apply some reasoning to establish causality remains to be seen.

Bumblebees are capable of not just learning from behaviour demonstrated by other bees, but they can spontaneously improve on that behaviour (Loukola, Perry, Coscos, & Chittka, 2017). Bumblebees can learn to roll balls to a specific location to achieve a sugar reward, and they learn faster if the behaviour is demonstrated. If the demonstrator shows moving a ball from a long distance away, bumblebees will spontaneously pick the ball that needs to be moved the shortest distance to the goal (Loukola et al., 2017). The extent to which this is causal reasoning is not clear, and neither is the nature of that reasoning.

The insect brain then gives us an example of very simple evolved dual-process cognition. The systems interact and through their interaction the problem space is refined by attending to what matters most such that innate behavioural responses can be modified or inhibited. The mushroom bodies are younger than the sensory lobes and pre-motor systems of the arthropod brain (Strausfeld et al., 1998), but we see mushroom bodies contributing to both Type 1-like and Type 2-like processes. This would argue against an evolutionary model of Type 1 processes being simple and coming first and Type 2 processes coming later, inviting further discussion of how dual-process systems might have evolved.

5. How dual processes of reasoning evolve

Dual processes of reasoning operate with overlapping domains on the same functional space (Osman, 2004). This meets the definition of a degenerate system (Maleszka et al., 2014), and we argue that this degenerate organisation is a vital aspect of a dual process system's robustness and flexibility.

Degenerate systems contain distinctively different elements that perform overlapping sets of functions (Maleszka et al., 2014). The concept of degeneracy is distinct from redundancy; a redundant system contains multiple elements with the same set of functions, whereas degenerate elements possess distinct but overlapping sets of functions (Edelman & Gally, 2001; Mason, 2010; Whitacre & Bender, 2010). Degeneracy is a common and important feature of the organisation of biological systems. It occurs in cognitive neuroanatomy (Friston & Price, 2003; Noppeney, Friston, & Price, 2004; Park & Friston, 2013), the genetic code (Luo, 1988; Subramaniam, Pan, & Cluzel, 2013), immune systems (Cohn, 2005), and genomic regulation (Maleszka et al., 2014).

In the immune system degeneracy of receptors means many different lymphocyte clones respond to any one specific antigen (Cohen, Hershberg, & Solomon, 2004). This accelerates both the speed and robustness of the immune response to a novel antigen, with multiple distinct elements acting in the same functional space (Cohen et al., 2004).

Degeneracy enhances the capacity of a system to adapt to a novel perturbation through either plasticity or evolutionary change which increases functional diversity (Cohen et al., 2004). If multiple elements support any one function then the operation of that function can be maintained even while a subset of elements are adapting over

phenotypic or evolutionary timescales. Because each element is distinct they can react to an environmental change in unique ways. As a consequence the system as a whole is able to more stably explore a greater diversity of possible states. Over time, degenerate elements of a system might diverge functionally through either the accumulation of mutations over evolutionary time, or phenotypic plasticity driving functional divergence. If they do so the total functional capacity of the system is expanded.

Anderson (2010) has argued that new cognitive functions evolve by a process of neural reuse, which proposes that neural circuits established for one purpose can be modified during evolution or development for different functions without losing their original functions. This process also depends on degenerate organisation for it to occur (Anderson, 2010; Mesulam, 1990). Each cognitive function is represented in multiple circuits with partially overlapping functional domains so that the mapping of anatomical substrate onto neural computation is both many to one and one to many (Anderson, 2010; Mesulam, 1990). This enables circuits to functionally diverge through plastic or evolutionary processes without compromising overall system performance in the process. Anderson (2010) argues that forms of autobiographical memory have reused motor systems. Linguistic processing has reused spatial and motor systems. Numeric processing has reused sensorimotor systems (Anderson, 2010). Neural reuse, enabled by degenerate organisation, has facilitated the evolution of new forms of cognition, including many forms that are important for Type 2 processes of reasoning such as linguistic and numeric processes.

Degeneracy crops up almost everywhere in nervous systems. The nematode worm's nervous system contains just 302 neurons. It is tremendously size constrained, but even in this system degeneracy is an organisational motif. For example, nematodes perform a random local search if they have exhausted a food patch. This behaviour is controlled by two parallel and structurally distinct chemo- and mechano-sensory circuits each using a similar glutamate receptor module. Either circuit alone is sufficient to trigger local search (López-Cruz et al., 2019). Simple nematode olfactory learning circuits also involve multiple parallel degenerate pathways linking chemo-receptors to command neurons via differing numbers of interneurons (Gray, Hill, & Bargmann, 2005). Any one of these pathways could functionally diverge over either physiological or evolutionary time to change its function without compromising the behavioural capacity of the organism.

In mammals, sensory processing has been found to occur via two parallel but interacting streams in vision (Goodale & Milner, 1992), audition (Hickok & Poeppel, 2007), and olfaction (Frasnelli et al., 2012). Similarly in insects it has been found that both olfactory (Brill et al., 2013) and vision (Li, Mahoney, Jacob, & Caron, 2020; Vogt et al., 2016; Yagi, Mabuchi, Mizunami, & Tanaka, 2016) utilise two parallel streams of processing. Motor control has been found to be governed by multiple parallel but interacting pathways in both mammals and insects (Barron, Gurney, Meah, Vasilaki, & Marshall, 2015; Namiki, Iwabuchi, Pansopha Kono, & Kanzaki, 2014) as well as the simple brains of nematode worms (López-Cruz et al., 2019). This suggests that brains across the animal kingdom routinely employ multiple degenerate pathways in their various calculations.

Computational models by Tononi, Sporns, and Edelman (1999) suggest degeneracy significantly increases evolvability while remaining robust across a wide variety of network types. Systems with high evolvability possessed elements that were functionally parallel but with high levels of interaction between them, and this motif was suggested to be a key characteristic allowing biological systems to remain robust despite being flexible and adaptable (Whitacre & Bender, 2010). Highly degenerate systems where different elements can affect output in a similar way whilst maintaining independence were seen to be the most evolvable (Tononi et al., 1999). Agent-based simulations by Whitacre and Bender (2010) again found that degeneracy significantly increased evolvability and robustness across a wide variety of network topologies and sizes. They modelled the relationship between agents and task

capabilities. Task completion was beyond the capacity of any single agent meaning agents had to cooperate to achieve task completion: different agents contributing different parts of the task. When the relationship between agents and task capabilities was highly degenerate task completion between highly robust against local disruptions or local resource limitations. In response to local resource limitations degeneracy can “transform local compensatory effects into long compensatory pathways” (Whitacre & Bender, 2010). They argue that degenerate systems could act generally as a catalyst for more open-ended evolution (Whitacre & Bender, 2010). These results suggest that increasing the degeneracy of reasoning would increase its evolvability, but we note that it makes no argument regarding how the functional space should be distributed across systems (i.e. what characteristics each process should possess). Benefits to robustness, flexibility, and evolvability apply to any degenerate organisation.

The different processing channels of the prime and non-prime lobes of the fly mushroom body discussed above are degenerate, and their degenerate relation to each other suggests how they might have evolved. Once channel could functionally diverge, while preserving the essential primitive function of the mushroom body of elementary learning.

The degeneracy argument favours models of dual systems which feature parallel but interacting processes (Tononi et al., 1999; Whitacre & Bender, 2010), as was found to be the best performing in simple computer simulations (Trimmer et al., 2008). This matches well with most dual processing theories which tend to posit a fundamental separation between the processes (see Table 1), in addition to the networks of AlphaGo where each system can function independently but only operate at the highest levels when interactions are allowed (Silver et al., 2016), and insects where each stream receives its own inputs and sends its own outputs yet feature some level of interaction (Aso, Hattori, et al., 2014; Takahashi et al., 2019). By operating in parallel it is ensured that each system acts on a distinct but overlapping portion of the functional space, rather than one being a subset of the other, making the combination of the two more robust and flexible.

Degenerate models might favour many distinct but overlapping elements rather than a single or dual system capable of operating on a spectrum between intuition and analytic thinking (Osman, 2004; Tononi et al., 1999; Whitacre & Bender, 2010). However, neural tissue is costly to build, maintain, and power (Sterling & Laughlin, 2015), putting an upper bound on the number of systems likely to evolve (Milli et al., 2017). Dual processes of reasoning combine a fast intuitive or associative system with a slower reflective or recursive system. Below we consider why dual processes of reasoning have evolved with these specific features.

6. Dual systems of reasoning and learning capacities

If degenerate systems are so common, are the examples discussed here in human cognition, artificial intelligence, and insect cognition fundamentally analogous to each other or separate examples of the benefits of a degenerate, multi-process system?

We would argue that the links between dual process theory and AIs like AlphaGo are substantial. These networks pair a fast neural network with a slower graphical system. As a rapid, feed forward associative system the neural network fits the definition of a Type 1 processes being intuitive and automatic (Osman, 2004; Silver et al., 2016). The graphical system of MuZero builds internal models of its world and uses this to predict future states, analysing the temporal dimension of the game to establish the effects of a move, similar to the rule-based calculations of Type 2 processes or the world-building of tripartite models (Osman, 2004; Stanovich, 2012; Schrittwieser et al., 2020).

AlphaGo is a degenerate system, with both of its two processes able to drive similar behaviour whilst functioning either cooperatively or independently (Silver et al., 2016; Tononi et al., 1999). Interactions between them prove reminiscent of dual process theories like Evans and Stanovich's, with rapid Type 1 responses being used to highlight select

stimuli and guide Type 2 processes onto those important or relevant states (Evans & Stanovich, 2013). Without this attentional filtering, Type 2 processes in AlphaGo significantly slowed and became increasingly inefficient (Silver et al., 2016), highlighting how each process not only supports but improves the other by refining and limiting the problem space. Interactions can also work the other way. In AlphaZero outputs from Type 2 processes generate more exploratory behaviour than its Type 1 processes would produce, significantly bolstering the stability of learning (Schrittwieser et al., 2020) and this matches with how learned behaviour can transition from Type 2 to Type 1 through practice in dual process theories (Osman, 2004). Models like MuZero also highlight the importance of world-building, with dedicated structures for understanding the rules of the game (Schrittwieser et al., 2020). This information can then be used to simulate future states and explore novel options, similar to tripartite models by Stanovich (2012).

While the similarities between human dual-process systems and current approaches to AI are clear, how the insect brain compares is more speculative. Behaviourally insects display learning patterns reminiscent of dual processes, with bumblebees initially solving problems using rules before switching to simple heuristics with practice (MaBouDi et al., 2020). Initial reactions and compulsive behaviours are governed by the prime lobes of the mushroom bodies while the non-prime lobes appear able to override these initial responses (Sun et al., 2017), reflecting Sloman's (Sloman, 1996) dual process theory like where Type 2 processes can suppress Type 1's initial reaction.

The insect brain highlights the importance of selective attention and its role in learning and decision making. With silenced mushroom bodies a fly's memory appears limited to simple snapshots, associating all observed stimuli with the punishment (Zhao et al., 2019). The presence of the mushroom bodies allowed the insect to selectively attend to the odour alone due to its repeated correlation with the punishment. Type 1 processes in insects are seen to respond to all incoming sensory signals (Strube-Bloss et al., 2011) and responds fast enough that it's plausible its processing is incorporated in other processes (Strube-Bloss et al., 2012).

While the insect examples do not match up perfectly with human formulations of dual processes of reasoning that emphasise intuition and reasoning, we find it compelling that we see a fast associative system engaged in rapidly analysing all input stimuli, and a slower, more deliberate and controlled system which examines selected representations in greater detail. The presence of these two systems is coupled with the emergence of selective attention, allowing for what is learned to be refined and independent of context. The insect example then gives a second example of interactions between dual processes refining a problem space to improve performance.

Attention can be thought of as not merely about focusing on a particular stimulus, but the selection of which stimulus should receive said focus. We would argue that in evolved brains dual process systems provide a solution of this attentional focus of cognitive resources. One system supports simple but rapid associative analysis of a large number of options. A second is another dedicated to slower but more nuanced analysis of the options of greatest interest, incorporating factors such as internal state and desired goals. By separating out different functions into separate systems each can specialise for its respective role.

The examples from both insets and AI emphasise how system can influence and improve the other, with Type 1 processes able to filter vast quantities of data for potential items of interest for Type 2, and Type 2 processes generating more exploratory and informative actions for Type 1 to learn from. With both channels able to drive similar behaviours based on distinct processing, the system as a whole becomes more flexible and robust while also allowing for an animal to respond to stimuli which provoke a consistent response without an overt attentional shift. Our comparative analysis suggest that dual processes of reasoning may have originated from these early selective attention systems. Hence the establishment of a fast intuitive system coupled with a slower reflective system.

The addition of attention represents a significant transformation in

learning capacity and decision making. This motif is robust yet flexible. Our AI examples operate very differently to solve different classes of problems, but they also show the motif of coupling a fast associative system with a reflective reasoning system to narrow down a problem space and focus cognitive resource. These we argue are the core advantages of dual processes of reasoning.

Author statement

MK and ABB conceived and wrote the paper.

Acknowledgements

This work was supported by an Australian Research Council Future Fellowship FT140100452 and grants TWCF0266 and TWCF0539 from the Templeton World Charity Foundation.

References

- Al-Anzi, B., & Zinn, K. (2018). Identification and characterization of mushroom body neurons that regulate fat storage in *Drosophila*. *Neural Development*, *13*(1). <https://doi.org/10.1186/s13064-018-0116-7>, 8.
- Amin, H., Apostolopoulou, A. A., Suárez-Grimalt, R., Vrontou, E., & Lin, A. C. (2020). Localized inhibition in the *Drosophila* mushroom body. *eLife*, *9*, 1–74. <https://doi.org/10.7554/ELIFE.56954>, 9.
- Anderson, M. L. (2010). Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences*, *33*, 245–313.
- Anthony, T., Tian, Z., & Barber, D. (2017). Thinking fast and slow with deep learning and tree search. *Advances in Neural Information Processing Systems, 2017-Decem*, 5361–5371, 5.
- Aptekar, J. W., Keleş, M. F., Lu, P. M., Zolotova, N. M., & Frye, M. A. (2015). Neurons forming optic glomeruli compute figure-ground discriminations in *Drosophila*. *Journal of Neuroscience*, *35*(19), 7587–7599. <https://doi.org/10.1523/jneurosci.0652-15.2015>, 5.
- Aso, Y., Hattori, D., Yu, Y., Johnston, R. M., Iyer, N. A., Ngo, T. B., ... Rubin, G. M. (2014). The neuronal architecture of the mushroom body provides a logic for associative learning. *eLife*, *3*. <https://doi.org/10.7554/eLife.04577>. e04577. 12.
- Aso, Y., Sitarman, D., Ichinose, T., Kaun, K. R., Vogt, K., Belliard-Guérin, G., ... Rubin, G. M. (2014). Mushroom body output neurons encode valence and guide memory-based action selection in *Drosophila*. *eLife*, *3*. <https://doi.org/10.7554/eLife.04580>. e04580. 12.
- Avargues-Weber, A., & Giurfa, M. (2013). Conceptual learning by miniature brains. *Proceedings of the Royal Society B*, *280*(1772). <https://doi.org/10.1098/rspb.2013.1907>
- Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: From ecological rationality to dual processes. *Behavioral and Brain Sciences*, *30*, 241–297.
- Barron, A. B., Gurney, K. N., Meah, L. F., Vasilaki, E., & Marshall, J. A. (2015). Decision-making and action selection in insects: Inspiration from vertebrate-based theories. *Frontiers in Behavioral Neuroscience*, *9*, 216. <https://doi.org/10.3389/fnbeh.2015.00216>, 8.
- de Bivort, B. L., & van Swinderen, B. (2016). Evidence for selective attention in the insect brain. *Current Opinion in Insect Science*, *15*, 9–15.
- Brembs, B. (2009). Mushroom bodies regulate habit formation in *Drosophila*. *Current Biology*, *19*(16), 1351–1355. <https://doi.org/10.1016/j.cub.2009.06.014>, 8.
- Brill, M. F., Rosenbaum, T., Reus, I., Kleineidam, C. J., Nawrot, M. P., & Rössler, W. (2013). Parallel processing via a dual olfactory pathway in the honeybee. *Journal of Neuroscience*, *33*(6), 2443–2456. <https://doi.org/10.1523/jneurosci.4268-12.2013>, 2.
- Carruthers, P. (2006). *The architecture of the mind*. Oxford: Clarendon Press.
- ten Cate, C., & Healy, S. D. (Eds.). (2017). *Avian cognition*. Cambridge U.K.: Cambridge University Press.
- Cohen, I. R., Hersherg, U., & Solomon, C. (2004). Antigen-receptor degeneracy and immunological paradigms. *Molecular Immunology*, *40*, 993–996.
- Cohn, M. (2005). Degeneracy, mimicry and crossreactivity in immune recognition. *Molecular Immunology*, *42*, 651–655.
- Coltheart, M. (1999). Modularity and cognition. *Trends in Cognitive Sciences*, *3*, 115–120.
- Cope, A. J., Vasilaki, E., Minors, D., Sabo, C., Marshall, J. A. R., & Barron, A. B. (2018). Abstract concept learning in a simple neural network inspired by the insect brain. *PLoS Computational Biology*, *14*(9). e1006435.
- Cross, F. R., Carvell, G. E., Jackson, R. R., & Grace, R. C. (2020). Arthropod intelligence? The case for Portia. *Frontiers in Psychology*, *11*, 2573. [https://doi.org/10.3389/fpsyg.2020.568049, 10.](https://doi.org/10.3389/fpsyg.2020.568049)
- De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition*, *106*, 1248–1299.
- Denes-Raj, V., & Epstein, S. (1994). Conflict between intuitive and rational processing: When people behave against their better judgment. *Journal of Personality and Social Psychology*, *66*, 819–829.
- Devaud, J. M., Papouin, T., Carcaud, J., Sandoz, J. C. B., .G., & Giurfa, M. (2015). Neural substrate for higher order learning in an insect: Mushroom bodies are necessary for configural discriminations. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, E5854.
- Edelman, G. M., & Gally, J. A. (2001). Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences USA*, *98*, 13763–13768.
- Erana, A. (2012). Dual process theories versus massive modularity hypotheses. *Philosophical Psychology*, *25*(6), 855–872. December 2012.
- Evans, J. S. B., & Over, D. E. (2013). *Rationality and reasoning*. Psychology/Eribaum (UK) Taylor & Francis. <https://doi.org/10.4324/9780203027677>
- Evans, J. S. B. T. (2019). Reflections on reflection: The nature and function of type 2 processes in dual-process theories of reasoning. *Thinking & Reasoning*, *25*, 383–415.
- Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, *8*(3), 223–241.
- Frasnelli, J., Lundström, J., Schöpf, V., Negroias, S., Hummel, T., & Lepore, F. (2012). Dual processing streams in chemosensory perception. *Frontiers in Human Neuroscience*, *6*. <https://doi.org/10.3389/fnhum.2012.00288>. OCTOBER 2012. 10.
- Friston, K., & Price, C. J. (2003). Degeneracy and redundancy in cognitive anatomy. *Trends in Cognitive Sciences*, *7*, 151–152.
- Giurfa, M., Zhang, S., Jenett, A., Menzel, R., & Srinivasan, M. V. (2001). The concepts of 'sameness' and 'difference' in an insect. *Nature*, *410*, 930–930.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*(1), 20–25. [https://doi.org/10.1016/0166-2236\(92\)90344-8, 1.](https://doi.org/10.1016/0166-2236(92)90344-8)
- Gray, J. M., Hill, J. J., & Bargmann, C. I. (2005). A circuit for navigation in *Caenorhabditis elegans*. *Proceedings of the National Academy of Science of the United States of America*, *102*, 3184–3191.
- Gronchi, G., & Giovannelli, F. (2018). Dual process theory of thought and default mode network: A possible neural foundation of fast thinking. *Frontiers in Psychology*, *9*, 1237.
- Groschner, L. N., Chan Wah Hak, L., Bogacz, R., DasGupta, S., & Miesenböck, G. (2018). Dendritic integration of sensory evidence in perceptual decision-making. *Cell*, *173*(4), 894–905. <https://doi.org/10.1016/j.cell.2018.03.075>, 5.
- Hammond, K. (1996). *Human judgement and social policy: Irreducible uncertainty, inevitable error* (Vol. xi). Oxford University Press. Retrieved from <https://philpapers.org/rec/RHAHAJ-2>.
- Hassabis, D., Kumaran, D., Summerfield, C., & Botvinick, M. (2018). Neuroscience-inspired artificial intelligence. *Neuron*, *95*(2), 245–258.
- Hattori, D., Aso, Y., Swartz, K. J., Rubin, G. M., Abbott, L. F., & Axel, R. (2017). Representations of novelty and familiarity in a mushroom body compartment. *Cell*, *169*(5), 956–969. <https://doi.org/10.1016/j.cell.2017.04.028>, 5.
- Hervé, J. C., & Derangeon, M. (2013). Gap-junction-mediated cell-to-cell communication. *Cell and Tissue Research*, *352*(1), 21–31. <https://doi.org/10.1007/s00441-012-1485-6>, 4.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393–402. <https://doi.org/10.1038/nrn2113>, 4.
- Hu, A., Zhang, W., & Wang, Z. (2010). Functional feedback from mushroom bodies to antennal lobes in the *Drosophila* olfactory pathway. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(22), 262–267. <https://doi.org/10.1073/pnas.0914912107>, 6.
- Inada, K., Tsuchimoto, Y., & Kazama, H. (2017). Origins of cell-type-specific olfactory processing in the *Drosophila* mushroom body circuit. *Neuron*, *95*(2), 357–367. <https://doi.org/10.1016/j.neuron.2017.06.039>, 7.
- Kahneman, D. (2011). *Thinking fast and slow*. Farrar, Straus and Giroux.
- Kamdayi, C., & Osvald, M. (2017). Ravens parallel great apes in flexible planning for tool use and bartering. *Science*, *357*, 202–204.
- Kamhi, J. F., Barron, A. B., & Narendra, A. (2020). Vertical lobes of the mushroom bodies are essential for view-based navigation in Australian Myrmecia ants. *Current Biology*, *30*(17), 3432–3437. <https://doi.org/10.1016/j.cub.2020.06.030>, 9.
- Krashes, M. J., DasGupta, S., Vreede, A., White, B., Armstrong, J. D., & Waddell, S. (2009). A neural circuit mechanism integrating motivational state with memory expression in *Drosophila*. *Cell*, *139*(2), 416–427. <https://doi.org/10.1016/j.cell.2009.08.035>, 10.
- Krashes, M. J., Keene, A. C., Leung, B., Armstrong, J. D., & Waddell, S. (2007). Sequential use of mushroom body neuron subsets during *Drosophila* odor memory processing. *Neuron*, *53*(1), 103–115. <https://doi.org/10.1016/j.neuron.2006.11.021>, 1.
- Li, J., Mahoney, B. D., Jacob, M. S., & Caron, S. J. C. (2020). Visual input into the *Drosophila* melanogaster mushroom body. *Cell Reports*, *32*(11), 108138. <https://doi.org/10.1016/j.celrep.2020.108138>, 9.
- Lin, A. C., Bygrave, A. M., De Calignon, A., Lee, T., & Miesenböck, G. (2014). Sparse, decorrelated odor coding in the mushroom body enhances learned odor discrimination. *Nature Neuroscience*, *17*(4), 559–568. <https://doi.org/10.1038/nn.3660>, 4.
- Lin, S., Oswald, D., Chandra, V., Talbot, C., Huetteroth, W., & Waddell, S. (2014). Neural correlates of water reward in thirsty *Drosophila*. *Nature Neuroscience*, *17*(11), 1536–1542. <https://doi.org/10.1038/nn.3827>, 11.
- Liu, L., Wolf, R., Ernst, R., & Heisenberg, M. (1999). Context generalization in *Drosophila* visual learning requires the mushroom bodies. *Nature*, *400*(6746), 753–756. <https://doi.org/10.1038/23456>, 8.
- Liu, Q., Yang, X., Tian, J., Gao, Z., Wang, M., Li, Y., & Guo, A. (2016). Gap junction networks in mushroom bodies participate in visual learning and memory in *Drosophila*. *eLife*, *5*. <https://doi.org/10.7554/eLife.13238>, 5.
- López-Cruz, A., Sordillo, A., Pokala, N., Liu, Q., McGrath, P. T., & Bargmann, C. I. (2019). Parallel multimodal circuits control an innate foraging behavior. *Neuron*, *102*, 407–419.
- Loukola, O. J., Perry, C. J., Coscos, L., & Chittka, L. (2017). Bumblebees show cognitive flexibility by improving on an observed complex behavior. *Science*, *355*(6327). <https://doi.org/10.1126/science.aag2360>, 2.

- Luo, L. F. (1988). The degeneracy rule of genetic code. *Origins of Life and Evolution of the Biosphere*, 18(1–2), 65–70.
- MaBouDi, H. D., Solvi, C., & Chittka, L. (2020). Bumblebees learn a relational rule but switch to a win-stay/lose-switch heuristic after extensive training. *Frontiers in Behavioral Neuroscience*, 14, 137. <https://doi.org/10.3389/fnbeh.2020.00137>, 8.
- Maleszka, R., Mason, P. H., & Barron, A. B. (2014). Epigenomics and the concept of degeneracy in biological systems. *Briefings in Functional Genomics*, 13(3), 191–202.
- Mason, P. H. (2010). Degeneracy at multiple levels of complexity. *Biological Theory*, 5, 277–288.
- Menzel, R., & Fischer, J. (Eds.). (2011). *Animal thinking*. Cambridge, Mass: MIT Press.
- Mesulam, M.-M. (1990). Large-scale neurocognitive networks and distributed processing for attention, language and memory. *Annals of Neurology*, 28, 597–613.
- Milli, S., Lieder, F., & Griffiths, T. L. (2017). When does bounded-optimal metareasoning favor few cognitive systems? In *Proceedings of the thirty-first AAAI conference on artificial intelligence* (pp. 4422–4428). December 2016.
- Moshman, D. (2000). Diversity in reasoning and rationality: Metacognitive and developmental considerations. *Behavioral and Brain Sciences*, 23(5), 689–690. <https://doi.org/10.1017/s0140525x00483433>, 10.
- Namiki, S., Iwabuchi, S., Pansopha Kono, P., & Kanzaki, R. (2014). Information flow through neural circuits for pheromone orientation. *Nature Communications*, 5(1), 5919. <https://doi.org/10.1038/ncomms6919>, 12.
- Newell, A. D. (1990). *Unified theories of cognition*. Harvard University Press. <https://doi.org/10.1002/wcs.1180>
- Newell, B. R., Lagnado, D. A., & Shanks, D. R. (2015). *Straight choices: The psychology of decision making*. New York: Psychology Press.
- Nityananda, V. (2016). Attention-like processes in insects. *Proceedings of the Royal Society B*, 283, 1986.
- Noppeney, U., Friston, K. J., & Price, C. J. (2004). Degenerate neuronal systems sustaining cognitive functions. *Journal of Anatomy*, 205, 433–442.
- Osman, M. (2004). An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin and Review*, 11(6), 988–1010. <https://doi.org/10.3758/BF03196730>
- Park, H., & Friston, K. (2013). Structural and functional brain networks: From connections to cognition. *Science*, 342, 579.
- Paulk, A. C., Stacey, J. A., Pearson, T. W., Taylor, G. J., Moore, R. J., Srinivasan, M. V., & Van Swinderen, B. (2014). Selective attention in the honeybee optic lobes precedes behavioral choices. *Proceedings of the National Academy of Sciences of the United States of America*, 111(13), 5006–5011. <https://doi.org/10.1073/pnas.1323297111>, 4.
- Perry, C. J., & Barron, A. B. (2013). Honey bees selectively avoid difficult choices. *Proceedings of the National Academy of Sciences*, 110(47), 19155–19159. <https://doi.org/10.1073/pnas.1314571110PNAS>
- Perry, C. J., Barron, A. B., & Cheng, K. (2013). Invertebrate learning and cognition: Relating phenomena to neural substrate. *WIREs Cognitive Science*, 4(5). <https://doi.org/10.1002/wcs.1248>
- Perry, C. J., Barron, A. B., & Chittka, L. (2017). The frontiers of insect cognition. *Current Opinion in Behavioral Sciences*, 16, 111–118. <https://doi.org/10.1016/j.cobeha.2017.05.011>
- Perry, C. J., & Chittka, L. (2019). How foresight might support the behavioral flexibility of arthropods. *Current Opinion in Neurobiology*, 54, 171–177.
- Schrittwieser, J., Antonoglou, I., Hubert, T., Simonyan, K., Sifre, L., Schmitt, S., ... Silver, D. (2020). Mastering Atari, Go, chess and shogi by planning with a learned model. *Nature*, 588(7839), 604–609. <https://doi.org/10.1038/s41586-020-03051-4>, 11.
- Sehdev, A., Mohammed, Y. G., Triphan, T., & Szyszka, P. (2019). Olfactory object recognition based on fine-scale stimulus timing in *Drosophila*. *iScience*, 13, 113–124.
- Shih, C. T., Sporns, O., Yuan, S. L., Su, T. S., Lin, Y. J., Chuang, C. C., ... Chiang, A. S. (2015). Connectomics-based analysis of information flow in the drosophila brain. *Current Biology*, 25(10), 1249–1258. <https://doi.org/10.1016/j.cub.2015.03.021>, 5.
- Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., ... Hassabis, D. (2016). Mastering the game of go with deep neural networks and tree search. *Nature*, 1–20.
- Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., ... Hassabis, D. (2018). A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science*, 362(6419), 1140–1144. <https://doi.org/10.1126/science.aar6404>, 12.
- Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., ... Hassabis, D. (2017). Mastering the game of Go without human knowledge. *Nature*, 550(7676), 354–359.
- Sitaraman, D., Aso, Y., Rubin, G. M., & Nitabach, M. N. (2015). Control of sleep by dopaminergic inputs to the drosophila mushroom body. *Frontiers in Neural Circuits*, 9 (November), 1–8. <https://doi.org/10.3389/fncir.2015.00073>
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119(1), 3–22.
- Solvi, C., Al-Khudairy, S. G., & Chittka, L. (2020). Bumble bees display cross-modal object recognition between visual and tactile senses. *Science*, 367(6480), 910–912. <https://doi.org/10.1126/science.aay8064>
- Stanovich, K. E. (2011). *Rationality and the reflective mind*. New York: Oxford University Press.
- Stanovich, K. E. (2012). Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory? In *In two minds: Dual processes and beyond* (pp. 55–89). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199230167.003.0003>, 3.
- Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, 23(5), 645–726. <https://doi.org/10.1017/S0140525X00003435>
- Sterling, P., & Laughlin, S. B. (2015). *Principles of neural design*. Cambridge MA: MIT Press.
- Strausfeld, N. J., Hansen, L., Li, Y., Gomez, R. S., & Ito, K. (1998). Evolution, discovery, and interpretations of arthropod mushroom bodies. *Learning and Memory*, 5(1–2), 11–37. <https://doi.org/10.1101/lm.5.1.11>, 5.
- Strube-Bloss, M. F., Herrera-Valdez, M. A., & Smith, B. H. (2012). Ensemble response in mushroom body output neurons of the honey bee outpaces spatiotemporal odor processing two synapses earlier in the antennal lobe. *PLoS ONE*, 7(11). <https://doi.org/10.1371/journal.pone.0050322>. e50322. 11.
- Strube-Bloss, M. F., Nawrot, M. P., & Menzel, R. (2011). Mushroom body output neurons encode odor-reward associations. *Journal of Neuroscience*, 31(8), 3129–3140. <https://doi.org/10.1523/jneurosci.2583-10.2011>, 2.
- Subramaniam, A. R., Pan, T., & Cluzel, P. (2013). Environmental perturbations lift the degeneracy of the genetic code to regulate protein levels in bacteria. *Proceedings of the National Academy of Sciences USA*, 110, 2419–2424.
- Sun, Y., Nern, A., Franconville, R., Dana, H., Schreier, E. R., Looger, L. L., ... Jayaraman, V. (2017). Neural signatures of dynamic stimulus selection in *Drosophila*. *Nature Neuroscience*, 20(8), 1104–1113. <https://doi.org/10.1038/nn.4581>, 8.
- van Swinderen, B., McCartney, A., Kauffman, S., Flores, K., Agrawal, K., Wagner, J., & Paulk, A. (2009). Shared visual attention and memory systems in the *Drosophila* brain. *PLoS One*, 4(6). <https://doi.org/10.1371/journal.pone.0005989>
- Takahashi, N., Nishino, H., Domae, M., & Mizunami, M. (2019). Separate but interactive parallel olfactory processing streams governed by different types of gabaergic feedback neurons in the mushroom body of a basal insect. *Journal of Neuroscience*, 39 (44), 8690–8704. <https://doi.org/10.1523/jneurosci.0088-19.2019>, 10.
- Thompson, V. A. (2013). Why it matters: The implications of autonomous processes for dual process theories—Commentary on Evans & Stanovich (2013). *Perspectives on Psychological Science*, 8, 253–256.
- Tibbetts, E. A., Agudelo, J., Pandit, S., & Riojas, J. (2019). Transitive inference in Polistes paper wasps. *Biology Letters*, 15, 20190015.
- Tomlin, D., Rand, D. G., Ludvig, E. A., & Cohen, J. D. (2015). The evolution and devolution of cognitive control: The costs of deliberation in a competitive world. *Scientific Reports*, 5, 11002.
- Tononi, G., Sporns, O., & Edelman, G. M. (1999). Measures of degeneracy and redundancy in biological networks. *Proceedings of the National Academy of Sciences of the United States of America*, 96(6), 3257–3262. <https://doi.org/10.1073/pnas.96.6.3257>
- Trimmer, P. C., Houston, A. I., Marshall, J. A., Bogacz, R., Paul, E. S., Mendl, M. T., & McNamara, J. M. (2008). Mammalian choices: Combining fast-but-inaccurate and slow-but-accurate decision-making systems. *Proceedings of the Royal Society B: Biological Sciences*, 275(1649), 2353–2361. <https://doi.org/10.1098/rspb.2008.0417>
- Tsao, C. H., Chen, C. C., Lin, C. H., Yang, H. Y., & Lin, S. (2018). *Drosophila* mushroom bodies integrate hunger and satiety signals to control innate food-seeking behavior. *eLife*, 7. <https://doi.org/10.7554/eLife.35264>, 3.
- Vogt, K., Aso, Y., Hige, T., Knapek, S., Ichinose, T., Friedrich, A. B., ... Tanimoto, H. (2016). Direct neural pathways convey distinct visual information to *drosophila* mushroom bodies. *eLife*, 5, 1–13. <https://doi.org/10.7554/eLife.14009>
- Wason, P. C., & Evans, J. S. B. T. (1974). Dual processes in reasoning? *Cognition*, 3(2), 141–154.
- Whitacre, J., & Bender, A. (2010). Degeneracy: A design principle for achieving robustness and evolvability. *Journal of Theoretical Biology*, 263(1), 143–153. <https://doi.org/10.1016/j.jtbi.2009.11.008>
- Wiederman, S. D., & O'Carroll, D. C. (2013). Selective attention in an insect visual neuron. *Current Biology*, 23(2), 156–161. <https://doi.org/10.1016/j.cub.2012.11.048>, 1.
- Wu, C. L., Shih, M. F. M., Lee, P. T., & Chiang, A. S. (2013). An octopamine-mushroom body circuit modulates the formation of anesthesia-resistant memory in *drosophila*. *Current Biology*, 23(23), 2346–2354. <https://doi.org/10.1016/j.cub.2013.09.056>, 12.
- Yagi, R., Mabuchi, Y., Mizunami, M., & Tanaka, N. K. (2016). Convergence of multimodal sensory pathways to the mushroom body calyx in *Drosophila melanogaster*. *Scientific Reports*, 6(1), 29481. <https://doi.org/10.1038/srep29481>, 9.
- Zhao, B., Sun, J., Zhang, X., Mo, H., Niu, Y., Li, Q., Zhong, & Y.. (2019). Long-term memory is formed immediately without the need for protein synthesis-dependent consolidation in *Drosophila*. *Nature Communications*, 10(1), 4550. <https://doi.org/10.1038/s41467-019-12436-7>