



**HAL**  
open science

# System-1 and System-2 realized within the Common Model of Cognition

Brendan Conway-Smith, Robert L. West

► **To cite this version:**

Brendan Conway-Smith, Robert L. West. System-1 and System-2 realized within the Common Model of Cognition. AAAI 2022 Fall Symposium, Association for the Advancement of Artificial Intelligence, Nov 2022, Arlington, Virginia, United States. 10.48550/arXiv.2305.09091 . hal-04104356

**HAL Id: hal-04104356**

**<https://hal.science/hal-04104356v1>**

Submitted on 24 May 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# System-1 and System-2 realized within the Common Model of Cognition

Brendan Conway-Smith<sup>1\*</sup>, Robert L. West<sup>1</sup>

<sup>1</sup>*Department of Cognitive Science, Carleton University, 1125 Colonel By Dr, Ottawa, ON, Canada*

## Abstract

Attempts to import dual-system descriptions of System-1 and System-2 into AI have been hindered by a lack of clarity over their distinction. We address this and other issues by situating System-1 and System-2 within the Common Model of Cognition. Results show that what are thought to be distinctive characteristics of System-1 and 2 instead form a spectrum of cognitive properties. The Common Model provides a comprehensive vision of the computational units involved in System-1 and System-2, their underlying mechanisms, and the implications for learning, metacognition, and emotion.

## Keywords

System 1, System 2, dual-process, cognitive modeling, common model, cognitive architecture

## 1.0. Introduction

Significant progress has been made in Artificial Intelligence (AI) by studying and attempting to copy human cognition. An important aspect of human intelligence is that humans tend to choose different types of cognitive processes to match the demands of a situation. The concept of System-1 and System-2 captures this idea by positing a dual-systems model where System-1 provides fast, heuristic-based thinking, and System-2 allows for slower, more rational thought. This maps onto folk psychological notions of rationality that contrast deliberate rational thought with fast impulsive thinking.

However, while the System-1 and System-2 dichotomy has provided a meaningful way to study different styles of thought, there is no agreed upon computational framework that integrates these two modes of thinking into a unified view of cognition. In this paper, we argue that the Common Model of Cognition, originally the ‘Standard Model’ [1] provides a unified framework for understanding System-1 and System-2. We will also show how recent criticisms of System-1 and System-2 can be clarified using the Common Model of Cognition to ground discussion.

## 2.0. System-1 and System-2

The terms System-1 and System-2 refer to a dual-system model that ascribes distinct characteristics to what are thought to be opposing aspects of cognition. Dual-systems theories (or dual-process theories) are common in psychology and posit that cognitive processes can be divided into two contrasting categories. A variety of theories have arisen due to different proposed ways of characterizing the division [2, 3, 4]. However, roughly speaking, dual-systems theories have divided thinking between the intuitive and the rational [5].

System-1 is considered to be evolutionarily old and characterized as fast, associative, emotional, automatic, and not requiring working memory. System-2 is considered to be more evolutionarily recent and characterized as slow, declarative, rational, effortful, and relying on working memory. System-1 and System-2 are often used in fields such as psychology, philosophy, neuroscience, and

---

*AAAI 2022 FALL SYMPOSIUM SERIES, Thinking Fast and Slow and Other Cognitive Theories in AI, November 17-19, Westin Arlington Gateway in Arlington, Virginia, USA*

✉ [brendan.conwaysmith@carleton.ca](mailto:brendan.conwaysmith@carleton.ca) (B. Conway-Smith); [robert.west@carleton.ca](mailto:robert.west@carleton.ca) (R. L. West)

ORCID [0000-0003-3936-6125](https://orcid.org/0000-0003-3936-6125) (B. Conway-Smith)



© 2022 Copyright for this paper by its authors.  
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).  
CEUR Workshop Proceedings (CEUR-WS.org)

artificial intelligence as a means for ontologizing the functional properties of human cognition. The neural correlates of System-1 and System-2 have also been examined [6].

Recently, however, this dual-system model has been criticized for lacking precision and conceptual clarity [7], leading to significant misconceptions [8] [9], and obscuring the dynamic complexities of psychological processes [10]. Much of this criticism stems from controversy over the alignment assumption. The alignment assumption refers to the claim that cognitive functions must align with either System-1 or System-2 [9]. From an AI perspective, the alignment assumption would be convenient, however, this assumption has been criticized as overly simplistic and some dual-systems theorists do not endorse it, referring instead to “typical correlates” rather than “defining features” [8].

Researchers in need of greater specificity have developed more detailed definitions of Systems-1 and 2. For example, Proust [11] has argued that a more precise computational definition is needed to understand the role of System-1 and 2 in metacognition (the use of higher level, or meta-level, processes to control cognition). Proust defined these systems in terms of their distinctive informational typologies, where System-1 metacognition is implicit, non-symbolic, and non-conceptual, while System-2 metacognition is explicit, symbolic, and conceptual.

Computational models of System-1 and System-2 have also been built. For example, Thomson et al. [12] argued that the expert use of heuristics (System-1) could be defined in terms of instance-based learning in ACT-R. In fact, there are numerous ways that cognitive models and cognitive architectures can and have been mapped onto the System-1 and 2 distinction. In some cases, the dual-process approach has been built directly into the architecture. For example, the CLARION architecture [13] and the LIDA architecture [14] have instantiated components that map directly onto characteristics of System-1 and System-2 type thinking.

However, while it is useful to work on modelling different aspects of System-1 and 2, the larger question is, in what sense is System-1 and System-2 a valid construct? What are the necessary and sufficient conditions that precisely define System-1 and 2? And what are the cognitive and neural alignments to System-1 and System-2?

## 2.1. Cognitive architectures

Evans [15], one of the originators of dual-system theory, has stated that an important issue for future research is the problem that “current theories are framed in general terms and are yet to be developed in terms of their specific computational architecture.”

Following Dennett [16], we argue that a computational description is essential for clarifying high-level, psychological characterizations, such as System-1 and System-2. At the time, Dennett received significant pushback on his view from psychology and philosophy, however, in our opinion, this was due to it being too early in the development of cognitive models to fully appreciate their value.

As Newell [17] noted, creating one-off computational models of individual psychological phenomena is also problematic, as it leads to a plethora of isolated “micro models”, unconstrained by considerations of forming an integrated agent architecture. Newell’s solution to this was the concept of a unified computational cognitive architecture. A cognitive architecture is a computationally implemented agent architecture that models the components of cognition and how they interact to produce psychological phenomena.

## 2.2. The Common Model

The Common Model of Cognition, originally the ‘Standard Model’ [1], is a consensus architecture that integrates decades of research on how human cognition functions computationally. The Common Model represents a convergence across cognitive architectures regarding the modules and components necessary for human-like intelligence. That is, the Common Model is a higher-level architectural specification that describes most, if not all, cognitive architectures capable of modeling the full range of human cognition. The Common Model has also been investigated by way of correlating its modules with their associated brain regions [18]. Neural imaging across a range of tasks strongly supports the Common Model as a leading candidate for modeling the functional organization of the human brain [19].

The Common Model has five components – working memory, perception, action, declarative memory, and procedural memory. Procedural memory is integral to how the Common Model operates, as it reacts and provides instruction to other modules based on working memory content. The simplest way of building procedural memory is as a production system acting on production rules (if-then rules). This is how it is implemented in ACT-R, yet not in other Common Model Architectures such as SOAR or Sigma. Since a production system is an easy way to describe and captures the essential action that we are interested in for this paper, we will describe procedural memory as a production system and procedural knowledge as productions. However, it is important to keep in mind that this is for communication purposes and that a production system represents the simplest possible mechanism for implementing procedural memory in the Common Model.

In terms of processing, the Common Model can be driven in two different ways. The first is through task specific productions. These are fast and automatic, in the sense that the system does not consider extra information – it knows how to proceed. The second way is through information stored in declarative memory. In this case, the necessary productions are not available, so information is requested from declarative memory, which is then used by productions to advance. While this maps onto System-1 and System-2, as we will see, the picture is more complex.

### 3.0. System-1

Researchers generally describe System-1 using a constellation of characteristics. Specifically, System-1 is described as fast, associative, emotional, automatic, and not requiring working memory [20, 21, 22]. System-1 is considered to be evolutionary old and present within other animals. It is composed of biologically programmed instinctive behaviors and operations that contain innate modules of the kind put forth by Fodor [23]. System-1 is not comprised of a single system but is an assembly of subsystems that are largely autonomous [24]. Automatic operations are usually described as involving minimal or no effort, and without a sense of voluntary control [20]. Researchers generally agree that System-1 is made of parallel and autonomous subsystems that output only their final product into consciousness (often as affect), which then influences human decision-making [15]. This is one reason the system has been called “intuitive” [25].

System-1 relies on automatic processes and shortcuts called heuristics - problem solving operations or rule of thumb strategies [26]. The nature of System-1 is often portrayed as non-symbolic, and has been associated with reinforcement learning [27] and neural networks [28]. Affect is integral to System-1 processes [29]. Affect based heuristics result from an individual evaluating a stimulus based on their likes and dislikes. In more complex decision-making, it occurs when a choice is either weighed as net positive (with more benefits than costs), or as net negative (less benefits than costs) [30].

System-1 can produce what are called “cognitive illusions” that can be harmful if left unchecked. For example, the “illusion of validity” is a cognitive bias where individuals overestimate their ability to accurately predict a data set, particularly when it shows a consistent pattern [31]. This is closely related to feelings-of-knowing that can provide both accurate and inaccurate epistemic signals [32]. Biases and errors within System-1 operate automatically and cannot be turned off at will. However, they can be offset by using System-2 to monitor System-1 and correct it.

### 3.1. System-1 in the Common Model

System-1 can be associated with the production system, which is the computational instantiation of procedural memory in the Common Model [33]. Procedural knowledge is represented as production rules (“productions”) which are modeled after computer program instructions in the form of condition-action pairings. They specify a condition that, when met, will perform a prescribed action. A production can also be thought of as an *if-then* rule. *If* it matches a condition, *then* it fires an action. Productions transform information to resolve problems or complete a task, and are responsible for state-changes within the system. Production rules fire automatically off of conditions in working memory [34]. They are considered to be automatic due to the fact that they are triggered without secondary evaluation. Neurologically, production rules correlate with the 50ms decision timing in the basal ganglia [35].

The Common Model production system has many of the properties associated with System-1 such as being fast, automatic, implicit, able to implement heuristics, and reinforcement learning. However, the Common Model declarative memory system also has some of the properties associated with System-1. Specifically, associative learning and the ability to implement heuristics that leverage associative learning [12]. Here, it is important to understand that the Common Model declarative memory cannot operate without the appropriate productions firing, and without the use of working memory. Therefore, from a Common Model perspective, System-1 minimally involves productions firing based on working memory conditions. However, it can also involve productions directing declarative memory retrieval, which also relies on working memory.

Based on this, System-1 cannot be defined as being uniquely aligned with either declarative or procedural memory. System-1 activity must necessarily involve production rules and working memory, and can also include declarative knowledge. A partial exception to this may be the direct links between perception and action. It is possible, in the CMC, to have responses that are directly triggered by perception. These connections, it turns out, are important for fitting to brain data [19].

## 4.0. System-2

Researchers generally understand System-2 in terms of a collection of cognitive properties characterized as slow, propositional, rational, effortful, and requiring working memory [4, 20, 35]. System-2 involves explicit propositional knowledge that is used to guide decision-making [37]. Propositional knowledge is associated with relational knowledge [38] which represents entities (e.g.: John and Mary), the relation between them (e.g.: loves) and the role of those entities in that relation (e.g.: John loves Mary). Higher level rationality in System-2 is also said to be epistemically committed to logical standards [6]. System-2 processes are associated with the subjective experiences of agency, choice, and effortful concentration [35]. The term “effortful” encompasses the intentional, conscious, and more strenuous use of knowledge in complex thinking. Higher level rationality is considered responsible for human-like reasoning, allowing for hypothetical thinking, long-range planning, and is correlated with measures of general intelligence [15].

Researchers have studied various ways in which System-2’s effortful processes can intervene in System-1 automatic operations [5]. Ordinarily, an individual does not need to invoke System-2 unless they notice that System-1 automaticity is insufficient or risky. System-2 can intervene when the anticipated System-1 output would infringe on explicit rules or potentially cause harm. For example, a scientist early in their experiment may notice that they are experiencing a feeling of certainty. System-2 can instruct them to resist jumping to conclusions and to gather more data. In this sense, System-2 can monitor System-1 and override it by applying conceptual rules.

## 4.1. System-2 in the Common Model

Laird [39] draws on Newell [40], Legg and Hutter [41] and others to equate rationality with intelligence, where “an agent uses its available knowledge to select the best action(s) to achieve its goal(s).” Newell’s Rationality Principle involves the assumption that problem-solving occurs in a problem space, where knowledge is used to navigate toward a desired end. As Newell puts it, “an agent will use the knowledge it has of its environment to achieve its goals” [42]. The prioritizing of knowledge in decision-making corresponds with the principles of classical computation involving symbol manipulation and transformation.

The Common Model architecture fundamentally distinguishes between declarative memory and procedural memory. This maps roughly onto the distinction between explicit and implicit knowledge; while declarative knowledge can be made explicitly available to working memory, procedural knowledge operates outside of working memory and is not directly accessible. However, declarative knowledge can also function in an implicit way. The presence of something within working memory does not necessarily mean it will be consciously accessed [43].

Higher level reasoning involves the retrieval of declarative information, representing propositional information, into working memory to assist in calculations and problem-solving operations. This appears to correlate with what System-2 researchers describe as “effortful” as this

requires more computational resources (i.e., more production cycles) to manage the flow of information through limited space in working memory. As Kahneman points out, System-1 can involve knowledge of simple processes such as  $2+2=4$ . However, more complex operations such as  $17 \times 16$  require calculations that are effortful, a characteristic that is considered distinctive of System-2 [20].

Effort, within the Common Model, involves greater computational resources being allocated toward a task. Moreover, the retrieval and processing of declarative knowledge requires more steps and more processing time when compared to the firing of productions alone. This longer retrieval and processing time can also account for the characteristic of “slow” associated with System-2.

## 5.0. Effort in System-1 and 2

The concept of “effort” makes up a significant and confusing dimension of System-1 and System-2. While it is mainly associated with System-2 rationality, a precise definition of “effort” remains elusive and is largely implicit in discussions of System-1 and System-2. Because System-2 is considered to have a low processing capacity, its operations are associated with greater effort and a de-prioritizing of irrelevant stimuli [44].

Effort can be associated with complex calculations in System-2 to the extent that it taxes working memory. Alternatively, effort can be associated with System-2’s capacity to overrule or suppress automatic processes in System-1 [20]. For example, various System-1 biases (such as the “belief bias”) can be subdued by instructing people to make a significant effort to reason deductively [45]. The application of formal rules to control cognitive processes is also called metacognition – the monitoring and control of cognition [46, 47]. Researchers have also interpreted metacognition through a System-1 and System-2 framework [48, 49]. System-1 metacognition is thought to be implicit, automatic, affect-driven, and not requiring working memory. System-2 metacognition is considered explicit, rule-based, and relying on working memory.

While the concept of “effort” is considered to be the monopoly of System-2, a computational approach suggests that effort is instead a continuum – with low effort cognitive phenomena being associated with System-1, and high effort cognitive phenomena being associated with System-2.

### 5.1. Effort in the Common Model

The Common Model helps to elucidate how “effort” can be present in System-1 type operations in the absence of other System-2 characteristics. While neither dual-system theories nor the Common Model contain a clear definition of “effort”, computational characteristics associated with effort can be necessitated by System-1. For instance, “effort” is often associated with the intense use of working memory. However, the Common Model requires working memory (along with its processing limitations) for both System-1 and System-2 type operations. There is little reason why System-1 should necessarily use less working memory than System-2 in the Common Model. Instead, it would depend on the task duration and intensity.

System-1 and System-2 can also be clarified by importing Proust’s [11] more precise account. Proust attempted to elucidate these two systems by claiming that they should be distinguished by their distinctive informational formats (System-1 non-conceptual; System-2 conceptual). In this sense, System-1 metacognition can exert effortful control while simultaneously being implicit and non-conceptual. For example, consider a tired graduate student attending a conference while struggling not to fall asleep. An example of System-1 metacognition would involve the context implicitly prompting them to feel nervous, noticing their own fatigue, and then attempting to stay awake. This effort is context-driven, implicit, non-conceptual, and effortful. Alternatively, System-2 metacognition can exert effort by way of explicit concepts, as in the case of a tired conference-attendee repeating the verbal instruction “try to focus”. Both scenarios could be modelled using the Common Model, and to reiterate, there is hardly any reason why System-1 should require less effort.

One way to think about effort is in terms of the expense of neural energy. In this sense, effort can be viewed as the result of greater caloric expenditure in neurons. The neural and computational dynamics responsible for the effortful control of internal states have shown to be sensitive to performance incentives [50]. Research also indicates that the allocation of effort as cognitive control

is dependent on whether a goal's reward outweighs its costs [51]. Both of these relate to reinforcement learning, which is associated with System-1. Additionally, effort in the Common Model can also be thought of in terms of the number of production cycles, where more cycles equate to more effort. In these terms, high effort is synonymous with slow processing, in that it takes longer to process a task. However, the number of production cycles would remain constant per unit of time.

Finally, it may be the case that overall effort is not the appropriate metric. Instead, System-2 thinking may be associated with a particular *type* of effort. Models of rational thinking in the Common Model architectures tend to use productions to manipulate information in working memory. For example, a production might take information from one chunk, modify it, and insert it into another chunk. This sort of read/write activity may in fact be what humans find effortful. Another possibility is that effort is related to the number of modules in use.

## 6.0. Emotion in System-1 and 2

Emotion and affect play a vital role in the distinction between System-1 and System-2 processes [20, 52]. Decisions in System-1 are largely motivated by an individual's implicit association of a stimulus with an emotion or affect (feelings that something is bad or good). Behavior that is motivated by emotion or affect is faster, more automatic, and less cognitively expensive. One evolutionary advantage of these processes is that they allow for split-second reactions that can be crucial for avoiding predators, catching food, and interacting with complex and uncertain environments.

Emotions can bias or overwhelm purely rational decision processes, but they can also be overridden by System-2 formal rules. While emotions and affect have historically been cast as the antithesis of reason, their importance in decision-making is being increasingly investigated by researchers who give affect a primary role in motivating decisions [53, 54]. Some maintain that rationality itself is not possible without emotion, as any instrumentally rational system must necessarily pursue desires [55].

## 6.1. Emotion in the Common Model

Feelings and emotions have strong effects on human performance and decision-making. However, there is considerable disagreement over what feelings and emotions are and how they can be incorporated into cognitive models. However, while philosophical explanations of emotion have been debated, numerous Common Model accounts of emotional phenomena have been created. Somatic markers have been modeled as emotional tags attached to units of information [56]. Low-level appraisals have been modeled as architectural self-reflections on factors such as expectedness, familiarity, and desirability [57]. Core affect theory has been modeled to allow agents to prioritize information using emotional valuation [58]. Alarm has been modeled as productions operating in the amygdala [59]. Affect and feelings have been modeled by treating them as non-propositional representations in working memory or "metadata" [60].

Overall, the question of how to model emotion and affect in the Common Model remains unresolved. However, a recent (2022) workshop on emotion in the Common Model has reached some level consensus:

1. Emotion acts in parallel to evaluate the threat levels and the desirability of both external states (e.g., an approaching tiger) and internal states (e.g., noticing you're fatigued and losing focus in a meeting).
2. Emotion can access information from working memory.
3. Emotion can input information into working memory.
4. Emotion can influence other modules through parameter adjustments (e.g., raise noise levels, lower thresholds, etc.).

From this viewpoint, emotion plays an important role in System-1 and System-2 thinking. One reason why System-2 and rational thinking have been thought of as separate from emotion is that rational thought requires sustained focus, which often means ignoring emotional distractions. However, from a Common Model perspective, the desire to attain a goal associated with the outcome of rational

thought is also emotional. This emotion arguably remains in the background as it is required by System-2 processes.

## 7.0. Learning in System-1 and 2

Researchers have associated System-1 with learning that is automatic, fast, and implicit, while System-2 learning is considered to be deliberate, slow, and explicit [36]. System-1 implicit learning usually involves subjects being unaware of what they are learning [61]. In contrast, System-2 explicit learning entails the intentional learning of information such as memorizing a list of word pairs. Research supports a distinction between implicit and explicit learning. Evidence shows that while subjects with amnesia have a reduced capacity for explicit learning, their implicit learning can remain [62]. The distinction between implicit and explicit learning is strongly related to learning in procedural memory and declarative memory, which are fundamental to psychology and neuroscience [63].

In the case of skill learning, Fitts & Posner [64] advanced a three-stage skill acquisition model whereby slow explicit knowledge, when repeatedly practiced, becomes converted into fast implicit knowledge. From a dual-system perspective, this is a process by which cognitive operations “migrate” from System-2 to System-1 as greater skill is developed [65].

## 7.1. Learning in the Common Model

As Laird and Mohan [66] noted, the Common Model can be thought of as having two levels of learning. Level 1 learning includes all automatic architectural learning mechanisms. This includes associative learning and episodic learning in declarative memory, as well as procedural compilation and reinforcement learning in procedural memory. Level 2 learning involves knowledge-based metacognitive strategies that create experiences for Level 1 mechanisms to learn from, such as deliberate practice and studying. Metacognitive strategies can be encoded into procedural memory for more efficient use of Level 1 mechanisms, but can also be encoded as declarative knowledge that is interpreted by procedural knowledge.

The Common Model also compiles stored instructions from declarative memory into procedural memory through repeated interpretation and practice [67]. When applied to Proust’s [11] dual-system framework for metacognition, this allows System-2 metacognition to be compiled into System-1 metacognition. It provides a source for the production-based metacognitive strategies discussed above (although they can also be learned through implicit, automatic mechanisms). This lends insight into the possible mechanisms that underlie metacognitive skill learning. Productions for higher-level logical thinking are also likely acquired this way (although some may be innate).

## 8.0. Conclusion

The following insights arise from grounding System-1 and System-2 in the Common Model:

1. *Both System-1 and System-2 rely on procedural memory.* While System-1 is more directly driven by the fast, automatic actions of the production system, System-2 is also reliant on the production system. Even when System-2 is driven primarily by explicit declarative knowledge, it requires the production system to retrieve and act on that knowledge.
2. *System-2 can involve emotion.* System-2 goal-directed rationality requires affect in (at least) the form of a preferred desired end. Further, according to the Common Model, there are multiple routes for System-2 rationality to be influenced by System-1 affective biases.
3. *Both System-1 and System-2 require working memory.* While the conventional view is that System-1 does not require working memory, the constraints of the Common Model necessitate it. Production rules (procedural knowledge) are activated by the content of working memory. Hence working memory is required for even the simplest System-1 processing.
4. *Effort is not well defined.* There is little evidence that System-2 requires more effort than System-1. Rather, humans seem to be more conscious of the effort required to maintain focused rational thought over longer periods of time.



Regardless of whether one adopts the Common Model architecture, researchers should be cautious of assuming that System-1 and System-2 can be treated as separate, dichotomous modules. The framework is far from agreed upon and deep issues continue to be unresolved.

Since Descartes, dualism has continually been reimagined as mind and soul, reason and emotions, and opposing modes of thought (e.g.: System-1 and System-2). These represent our attempts to make sense of our own minds, its processes, and how this understanding maps onto our personal experiences. Clearly, System-1 and System-2 capture something deeply intuitive about the phenomenology of cognition.

Interpreting System-1 and System-2 within the Common Model results in our concluding that the “alignment assumption” (that the two systems are opposites) is a false dichotomy. There are, of course, cases where all properties of System-1 and System-2 are cleanly bifurcated on either side. However, between these two extremities lies a spectrum where the characteristics are mixed. In fact, from a Common Model perspective, it would be more accurate to say that the relationship is hierarchical, with System-2 built on top of System-1. In this sense, a “levels” characterization might be more appropriate. A good candidate for this is Newell’s distinction between the Cognitive Level and the Knowledge Level [40], which conceptually maps onto System-1 and System-2.

Interpreting System-1 and System-2 within the Common Model suggests that System-2 is emergent from the dynamic interactions within System-1, *in humans*. This raises the question of whether this is the best way to build System-1 and System-2 in an AI. Possibly, building them as separate systems would avoid certain shortcomings of the human system. Certainly, it can be more desirable from a design and tractability perspective. However, it is also possible that a two systems approach would be less flexible and less adaptive than an emergent levels approach.

The Common Model of Cognition provides a comprehensive view of the computational units involved in System-1 and System-2 type processes. By grounding dual-process models within the framework of the Common Model we gain a clearer understanding of the underlying mechanism involved.

## References

- [1] J. E., Laird, C. Lebiere, P.S. Rosenbloom, A standard model of the mind: Toward a common computational framework across artificial intelligence, cognitive science, neuroscience, and robotics." *Ai Magazine* 38, no. 4 (2017): 13-26. <https://doi.org/10.1609/aimag.v38i4.2744>
- [2] P. C. Wason, J. S. B. Evans, Dual processes in reasoning?. *Cognition* 3, no. 2 (1974): 141-154.
- [3] K.E. Stanovich, Who is rational?: Studies of individual differences in reasoning. Psychology Press, 1999. <https://doi.org/10.4324/9781410603432>
- [4] F. Strack, R. Deutsch, Reflective and impulsive determinants of social behavior. *Personality and social psychology review* 8, no. 3 (2004): 220-247.
- [5] D. Kahneman, A perspective on judgment and choice: mapping bounded rationality. *American psychologist* 58, no. 9 (2003): 697. <https://doi.org/10.1037/0003-066X.58.9.697>
- [6] T. Tsujii, S. Watanabe, Neural correlates of dual-task effect on belief-bias syllogistic reasoning: a near-infrared spectroscopy study. *Brain research* 1287 (2009): 118-125.
- [7] G. Keren, Y. Schul, Two is not always better than one: A critical evaluation of two-system theories. *Perspectives on psychological science* 4, no. 6 (2009): 533-550.
- [8] G. Pennycook, W. De Neys, J.S.B. Evans, K.E. Stanovich, V.A. Thompson, The mythical dual-process typology. *Trends in Cognitive Sciences* 22, no. 8 (2018): 667-668.
- [9] J. De Houwer, Moving beyond System 1 and System 2: Conditioning, implicit evaluation, and habitual responding might be mediated by relational knowledge. *Experimental Psychology* 66, no. 4 (2019): 257. <https://doi.org/10.1027/1618-3169/a000450>
- [10] A. Moors, Automaticity: Componential, causal, and mechanistic explanations. *Annual review of psychology* 67, no. 1 (2016): 263-287.
- [11] J. Proust, *The philosophy of metacognition: Mental agency and self-awareness*. Oxford, 2013.
- [12] R. Thomson, C. Lebiere, J.R. Anderson, J. Staszewski, A general instance-based learning framework for studying intuitive decision-making in a cognitive architecture. *Journal of Applied Research in Memory and Cognition* 4, no. 3 (2015): 180-190.
- [13] R. Sun, S. Paul, T. Chris, The interaction of the explicit and the implicit in skill learning: a dual-process approach." *Psychological review* 112, no. 1 (2005): 159.
- [14] U. Faghihi, C. Estey, R. McCall, S. Franklin, A cognitive model fleshes out Kahneman's fast and slow systems. *Biologically Inspired Cognitive Architectures* 11 (2015): 38-52.
- [15] J.S.B. Evans, In two minds: dual-process accounts of reasoning. *Trends in cognitive sciences* 7, no. 10 (2003): 454-459. <https://doi.org/10.1016/j.tics.2003.08.012>
- [16] D.C., Dennett, *Brainstorms: Philosophical essays on mind and psychology*. MIT press (2017).
- [17] A. Newell, You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. (1973).
- [18] Z. Steine-Hanson, N. Koh, A. Stocco, Refining the common model of cognition through large neuroscience data. *Procedia computer science* 145 (2018): 813-820.
- [19] A. Stocco, C. Sibert, Z. Steine-Hanson, N. Koh, J.E. Laird, C.J. Lebiere, P. Rosenbloom, Analysis of the human connectome data supports the notion of a "Common Model of Cognition" for human and human-like intelligence across domains." *NeuroImage* 235 (2021): 118035.
- [20] D. Kahneman, *Thinking, fast and slow*. Macmillan (2011).
- [21] J. S. B. Evans, K.E. Stanovich, Dual-process theories of higher cognition: Advancing the debate. *Perspectives on psychological science* 8, no. 3 (2013): 223-241.
- [22] F. Strack, R. Deutsch, Reflective and impulsive determinants of social behavior. *Personality and social psychology review* 8, no. 3 (2004).
- [23] J. Fodor, *The modularity of mind*. Scranton. (1983).
- [24] K. E. Stanovich, R. F. West, Individual differences in reasoning: Implications for the rationality debate?. *Behavioral and brain sciences* 23, no. 5 (2000): 645-665.
- [25] D. Kahneman, A perspective on judgment and choice: mapping bounded rationality. *American psychologist* 58, no. 9 (2003): 697. <https://doi.org/10.1037/0003-066X.58.9.697>
- [26] H. A. Simon, Herbert, A behavioral model of rational choice." *The quarterly journal of economics* 69, no. 1 (1955): 99-118. doi:10.2307/1884852

- [27] A. G. Barto, S. S. Richard, S. B. Peter, Associative search network: A reinforcement learning associative memory. *Biological cybernetics* 40, no. 3 (1981): 201-211.
- [28] P. McLeod, K. Plunkett, E. T. Rolls, Introduction to connectionist modelling of cognitive processes. Oxford University Press (1998).
- [29] D. G. Mitchell, The nexus between decision making and emotion regulation: a review of convergent neurocognitive substrates. *Behavioural brain research* 217, no. 1 (2011): 215-231.
- [30] P. Slovic, M.L. Finucane, E. Peters, D.G. MacGregor, Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk and rationality. In *The feeling of risk*, pp. 21-36. Routledge, (2013). doi:10.1111/j.0272-4332.2004.00433.x
- [31] D. Kahneman, A. Tversky, On the psychology of prediction. *Psychological review* 80, no. 4 (1973): 237. <https://doi.org/10.1037/h0034747>
- [32] J. T. Hart, Julian, Memory and the feeling-of-knowing experience. *Journal of educational psychology* 56, no. 4 (1965): 208.
- [33] M. K. Singley, J. R. Anderson, The transfer of cognitive skill. No. 9. Harvard Uni. Press, 1989.
- [34] J. R. Anderson, Knowledge representation. *Rules of the mind* (1993): 17-44.
- [35] A. Stocco, C. Lebiere, J. R. Anderson, Conditional routing of information to the cortex: A model of the basal ganglia's role in cognitive coordination. *Psychological review* 117, no. 2 (2010): 541.
- [36] K. Frankish, Dual-process and dual-system theories of reasoning. *Philosophy Compass* 5, no. 10 (2010): 914-926. <https://doi.org/10.1111/j.1747-9991.2010.00330.x>
- [37] S. Epstein, R. Pacini, Some basic issues regarding dual-process theories from the perspective of cognitive-experiential self-theory. (1999).
- [38] G. S. Halford, W. H. Wilson, S. Phillips, Relational knowledge: The foundation of higher cognition. *Trends in cognitive sciences* 14, no. 11 (2010): 497-505.
- [39] J. Laird, Intelligence, knowledge & human-like intelligence. *Journal of Artificial General Intelligence* 11, no. 2 (2020): 41-44. 10.2478/jagi-2020-0003
- [40] A. Newell, Unified theories of cognition, Harvard University Press, Cambridge, MA (1990).
- [41] S. Legg, M. Hutter, Universal intelligence: A definition of machine intelligence. *Minds and machines* 17, no. 4 (2007): 391-444.
- [42] A. Newell, The knowledge level. *Artificial intelligence* 18, no. 1 (1982): 87-127.
- [43] D. Wallach, C. Lebiere, Implicit and explicit learning in a unified architecture of cognition. *Attention and implicit learning* (2003): 215-250.
- [44] K. E. Stanovich, Who is rational?: Studies of individual differences in reasoning. Psychology Press (1999). <https://doi.org/10.4324/9781410603432>
- [45] J. S. B. Evans, T. Julie, L. Barston, P. Pollard, On the conflict between logic and belief in syllogistic reasoning. *Memory & cognition* 11, no. 3 (1983): 295-306.
- [46] J. H. Flavell, Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American psychologist* 34, no. 10 (1979): 906.
- [47] L. Fletcher, P. Carruthers, Metacognition and reasoning. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367, no. 1594 (2012): 1366-1378.
- [48] S. Arango-Muñoz, Two levels of metacognition. *Philosophia* 39, no. 1 (2011): 71-82.
- [49] N. Shea, A. Boldt, D. Bang, N. Yeung, C. Heyes, C.D. Frith, Supra-personal cognitive control and metacognition. *Trends in cognitive sciences* 18. (2014):186-193.
- [50] S. W. Egger, E. D. Remington, C. J. Chang, M. Jazayeri, Internal models of sensorimotor integration regulate cortical dynamics. *Nature neuroscience* 22, no. 11 (2019): 1871-1882.
- [51] A. Shenhav, S. Musslick, F. Lieder, W. Kool, T. L. Griffiths, J.D. Cohen, M. M. Botvinick, Toward a rational and mechanistic account of mental effort. *Annual review of neuroscience* 40, no. 1 (2017): 99-124. <https://doi.org/10.1146/annurev-neuro-072116-031526>
- [52] S. Chaiken, and Y. Trope, Dual-process theories in social psychology. Guilford Press, 1999.
- [53] R. B. Zajonc, Feeling and thinking: Preferences need no inferences. *American psychologist* 35, no. 2 (1980): 151. <https://doi.org/10.1037/0003-066X.35.2.151>
- [54] L. F. Barrett, P. Salovey, The wisdom in feeling: Psychological processes in emotional intelligence. Guilford Press (2002).
- [55] J. S. B. Evans, Spot the difference: distinguishing between two kinds of processing. *Mind & Society* 11, no. 1 (2012): 121-131.

- [56] A. R. Damasio, Descartes' error and the future of human life. *Scientific American* 271, no. 4 (1994): 144-144.
- [57] P. S. Rosenbloom, J. Gratch, V. Ustun, Towards emotion in sigma: from appraisal to attention. In *International conference on artificial general intelligence*, pp. 142-151. Springer, Cham (2015).
- [58] I. Juvina, O. Larue, A. Hough, Modeling valuation and core affect in a cognitive architecture: The impact of valence and arousal on memory and decision-making. *Cognitive Systems Research* 48 (2018): 4-24. <https://doi.org/10.1016/j.cogsys.2017.06.002>
- [59] R. L. West, J. T. Young, Proposal to add emotion to the standard model. In *2017 AAAI Fall Symposium Series* (2017).
- [60] R. L. West, B. Conway-Smith, Put Feeling into Cognitive Models: A Computational Theory of Feeling. In *Proceedings of ICCM 2019 17th International Conference on Cognitive Modelling*, (2019) [iccm-conference.neocities.org/2019/proceedings/papers/ICCM2019\\_paper\\_47.pdf](http://iccm-conference.neocities.org/2019/proceedings/papers/ICCM2019_paper_47.pdf)
- [61] P. A. Frensch, D. Runger, Implicit learning. *Current directions in psychological science* 12, no. 1 (2003): 13-18. <https://doi.org/10.1111/1467-8721.01213>
- [62] B. Milner, Les troubles de la memoire accompagnant des lesions hippocampiques bilaterales. *Physiologie de l'hippocampe* 107 (1962): 257-272.
- [63] L. R. Squire, Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. *Journal of cognitive neuroscience* 4, no. 3 (1992): 232-243.
- [64] P. M. Fitts, M. I. Posner, *Human performance*. Brooks/Cole, Belmont, CA 5 (1967): 7-16
- [65] D. Kahneman, S. Frederick, Attribute substitution in intuitive judgment. *Models of a man: Essays in memory of Herbert A. Simon* (2004): 411-432.
- [66] J. Laird, S. Mohan, Learning fast and slow: Levels of learning in general autonomous intelligent agents. In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1. (2018).
- [67] J. R. Anderson, Acquisition of cognitive skill. *Psychological review* 89, no. 4 (1982): 369.