Systems and Cognitive Neuroscience
(NEU 502A / PSY 502A / MOL 502A)

Spring 2025

Professor: Jonathan Cohen (jdc@princeton.edu)

AI: Alex Ku (alexku@princeton.edu)

Course Description

A survey of experimental & theoretical approaches to understanding how cognition arises in the brain. This complements NEU 501A, focusing on the mechanisms responsible for perception, attention, decision making, memory, cognitive & motor control, and planning, with emphasis on the representations involved & their transformations in the service of cognitive function. Source material will span neuroscience, cognitive science, & work on artificial systems. Relevance to neurodegenerative and neuropsychiatric disorders will also be discussed. This is the 2nd term of a double-credit core lecture course required of all Neuroscience Ph.D. students.

Course Structure

Lectures: M/Th 2-4:30

Format: the two classes each week will be divided into three parts: "orienting" lectures, a "deep dive" by a visiting lecturer, and a student paper presentation. There will be two orienting lectures by the course instructor (one each class); on Mondays, that will be followed by a deep dive lecture by another faculty member in PNI, and on Thursdays it will be followed by a student paper presentation.

Grading

Class participation: 50%
Paper presentation: 50%

Lectures and Readings

Note: The readings listed under each lecture are meant to be a resource, for additional background and/or deeper coverage of the material presented in the lectures. Asterisked readings are required.

Schedule

Week 1 ————————————————

SECTION 1: PERCEPTION, CONSTRAINT SATISFACTION and ASSOCIATIVE LEARNING

Class 1 (Monday Jan 27) — **Overview**

• Introduction and History (Cohen)

*__Symbolic vs. Connectionist Models:__ Minsky, M. & Papert, S. (1988), *Perceptrons*. Cambridge, MA: MIT Press:

*Prologue, pp. vii-xv
Epilogue, pp. 247-281

***Parallel Distributed Processing (PDP)**. McClelland, J. and Rumelhart, D. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 1: Foundations*. Cambridge, MA: MIT Press:

* Chapter 1. The appeal of parallel distributed processing, pp. 3-44
* Chapter 4. The PDP Framework for Information Processing, pp. 110-137
Chapter 2. A General Framework for Parallel Distributed Processing, pp. 45-76
Chapter 9. An introduction to linear algebra in parallel distributed processing, pp. 365-421

Spreading Activation Models. Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82, 407-428.

100 Step Challenge. Feldman, J. A., & Ballard, D. H. (1982). Connectionist models and their properties. *Cognitive Science*, *6*(3), 205-254.

Modularity and Generativity. Fodor, J. A. & Pylyshyn, Z. W. (1988) Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28, 3-71

Modeling. McClelland, J. L. (2009). The place of modeling in cognitive science. *Topics in Cognitive Science*, *1*(1), 11-38.

Class 2 (Thursday Jan 30) **— Perception**

- Perceptual Phenomena, Constraint Satisfaction and Attractor Networks (Cohen)
- Retina and V1 **(**Pillow**)**

[LAB: Necker Cube]

* **Attractor networks. Hopfield, J. J. (1982).** Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, *79*(8), 2554-2558.

* **The physiology of the retina.** Barlow, HB (1982). The senses, pp 102--113. Cambridge University Press.

* **Receptive fields of single neurones in the cat's striate cortex.** Hubel, DH & Wiesel, TN (1959). The Journal of Physiology, pp. 148, 574-591.

Perceptual Bistability. Gigante G, Mattia M, Braun J, Del Giudice P (2009) Bistable Perception Modeled as Competing Stochastic Integrations at Two Levels. *PLoS Computation Biology 5*(7): e1000430 [https://doi.org/10.1371/journal.pcbi.1000430]

Possible principles underlying the transformation of sensory messages. Barlow, H (1961). Sensory Communication, Cambridge, MA: MIT Press. pp. 217-234.

Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Olshausen BA & Field DJ (1996). *Nature*, 1996, 381, 607-609.

Week 2 ————————————

Class 3 (Monday Feb 3) — **Associative Learning and Topography in Cortex**

[LAB: Hebbian learning]

- Associative Learning and Feature Maps (Cohen)
- Topographic Organization of Visual Cortex (Graziano)

* **Semantic feature organization**. Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological cybernetics*, *43*(1), 59-69.

* **Cortical Organization of representations**. Graziano, M. S., & Aflalo, T. N. (2007). Mapping behavioral repertoire onto the cortex. *Neuron, 56*(2), 239-251.

Occular Dominance Columns. Miller, K. D., Keller, J. B., & Stryker, M. P. (1989). Ocular dominance column development: Analysis and simulation. *Science, 245*, 605-615.

* **Hebbian Learning**. Hebb, D. O. (1949). *The Organization of Behavior.* Introduction, xi–xix; and Chapter 4: The first stage of perception: growth of an assembly, 60–78.

Biological Plausibility. Song, S., Miller, K. D., & Abbott, L. F. (2000). Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nature neuroscience*, *3*(9), 919.

Class 4 (Thursday Feb 6) — **High Level Vision**

- Object Recognition and Faces (Gomez)

Cell-specific object representations. Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience*, *4*(8), 2051-2062.

"**Grandmother cells."** Gross, C. G. (2002). Genealogy of the "grandmother cell". *The Neuroscientist*, *8*(5), 512-518.

"Fusiform Face Area" (FFA). Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1476), 2109-2128.

Distributed object representations. Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in cognitive sciences*, *4*(6), 223-233.

Similarity and categorization in vision. Battleday, R. M., Peterson, J. C., & Griffiths, T. L. (2021). From convolutional neural networks to models of higher-level cognition (and back again). *Annals of the New York Academy of Sciences, 1505*(1), 55-78.

First connectionist model of neuronal response characteristics. Andersen, R. A., & Zipser, D. (1988). The role of the posterior parietal cortex in coordinate transformations for visual–motor integration. *Canadian journal of physiology and pharmacology*, *66*(4), 488-501.

Convolutional neural networks. LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.

Using deep learning to understand the brain. *Yamins, D. L., & DiCarlo, J. J. (2016). Using goal-driven deep learning models to understand sensory cortex. Nature Neuroscience, 19(3), 356-365.*

Circuits in deep learning. Olah, C., Cammarata, N., Schubert, L., Goh, G., Petrov, M., & Carter, S. (2020). Zoom in: An introduction to circuits. Distill, 5(3), e00024-001.

Week 3 ___________________________

SECTION 2: DECISION MAKING - INTEGRATION

Class 5 (Monday Feb 10) — **Dynamics of integration and decision making**

[LAB: Drift Diffusion Model]

- Accumulation to bounds models of Decision Making (Cohen)
- Neural Mechanisms of Perceptual Decision Making (Brody)

* **Integration Models of Decision Making**. Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*(4), pp.700-720.

* **Shadlen, M. N., & Newsome, W. T. (2001)**. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of neurophysiology*, *86*(4), 1916-1936

The Drift Diffusion Model. Ratcliff, R. (1978). A theory of memory retrieval. *Psychological review*, *85*(2), 59.

Brainwide analysis of decision making. *BondyAG, Charlton, JA, Luo TZ, Kopec CD, Stagnaro WM, Venditto SJC, Lynch L, Janarthanan S, Oline SN, Harris TD & Brody CD (2024). bioRxiv, 2024-08.*

Class 6 (Thursday Feb 13) — Optimization of decision making

- Optimization and Control of Decision Making (Cohen)

* **Reward rate optimization of decision making**. Cohen JD & Holmes P (2014). Optimality and some of its discontents. *Topics in Cognitive Science, 6*, 258-278

Reward rate optimization of decision making. Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review, 113(4)*, pp.720-738.

SECTION 3: REINFORCEMENT LEARNING - REWARD and NEUROMODULATION

Class 7 (Monday Feb 17) — **Reinforcement Learning**

[LAB: Reinforcement Learning]

- Reinforcement Learning (Cohen)
- Model-Free vs. Model-Based Learning, Dopamine and Basal Ganglia (Daw)

* **Reward systems.** Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annual Review of Psychology*, *57*(1), 87-115.

* **Dopamine and reward prediction**. Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience*, 16, 1936-1947.

* **Model-Free vs. Model-Based RL**. Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*(12), 1704.

Model-Free vs. Model-Based RL. Drummond, N., & Niv, Y. (2020). Model-based decision making and model-free learning. *Current Biology*, *30*(15), R860-R865.

Temporal difference learning. Sutton, R. S. & Barto, A. G. (1981) Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review*, 88, pages 135-170.

Hierarchical RL. Botvinick, M. M., Niv, Y., & Barto, A. C. (2009). Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. *Cognition*, *113*(3), 262-280.

Class 8 (Thursday Feb 20) — **Explore/exploit and noradrenergic neuromodulation**

- Explore / Exploit and the Locus Coeruleus-Norepinephrine System (Cohen)

* **Overview.** Cohen, J. D., McClure, S. M., & Angela, J. Y. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 933-942.

* **Adaptive Gain Hypothesis**. Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403-450.

Behavior and Brain Imaging. Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876.

Uncertainty and Exploration. Yu A, & Dayan P (2005). Uncertainty, neuromodulation, and attention. *Neuron*, *46*(4), 681-692.

Dopamine-norepinephrine interactions in exploration and learning. *McClure S, Gilzenrat M and Cohen J (2005). An exploration-exploitation model based on norepinephrine and dopamine activity. Advances in Neural Information Processing Systems, 18.*

Biologically-informed model of LC. Usher M, Cohen JD, Rajkowski J, Kubiak P & Aston-Jones G (1999). The role of locus coeruleus in the regulation of cognitive performance. *Science, 283*, 549-554.

Week 5 ————————————

SECTION 4: SEMANTIC MEMORY - STATISTICAL LEARNING and DISTRIBUTED REPRESENTATION

Class 9 (Monday Feb 24) — **Distributed Representation, Learning and Semantics**

[LAB: Rumelhart Network / Backpropagation Learning Algorithm]

- Distributed Representations and Semantics (Cohen)
- Backpropagation Learning, Internal Representations and Semantics (Cohen)

* **Model of Distributed Memory.** McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, *114*(2), 159.

* **Semantic representations**. Rumelhart, D. E., & Todd, P. M. (1993). Learning and connectionist representations. *Attention and Performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience*, 3-30.

* **Semantic representations and coherent covariation.** McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience*, 4(4), 310-322.

Family tree model. Hinton, G. E. (1986, August). Learning distributed representations of concepts. In *Proceedings of the eighth annual conference of the Cognitive Science Society* (Vol. 1, p. 12).

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of verbal learning and verbal behavior*, *8*(2), 240-247.

Backpropagation. Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. N*ature*, *323*(6088), 533.

Nettalk. Sejnowski, T. J., & Rosenberg, C. R. (1988). NETtalk: A parallel network that learns to read aloud. Neurocomputing: foundations of research.

Dimensional Structure and Learning. *Saxe, A. M., McClelland, J. L., & Ganguli, S. (2019). A mathematical theory of semantic development in deep neural networks. Proceedings of the National Academy of Sciences, 116(23), 11537-11546.*

Biological Plausibility. *O'Reilly, R. C. (2001). Generalization in interactive networks: The benefits of inhibitory competition and Hebbian learning. Neural Computation, 13(6), 1199-1241.*

Biological Plausibility. Xie, X., & Seung, H. S. (2003). Equivalence of backpropagation and contrastive Hebbian learning in a layered network. *Neural computation*, *15*(2), 441-454.

Biological Plausibility. Bozkurt, B., Pehlevan, C., & Erdogan, A. (2024). Correlative information maximization: a biologically plausible approach to supervised deep neural networks without weight symmetry. Advances in Neural Information Processing Systems, 36.

Deep Learning Ref - LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436.

Class 10 (Thursday Feb 27) — Bayesian Inference and Neural Networks

- Bayes in the Age of Intelligent Machines (Griffiths)

  * **Bayesian models of cognition.** L Griffiths, T., Kemp, C., & B Tenenbaum, J. (2008). *Bayesian models of cognition.*

  * **Bayesian inference and neural networks.** Griffiths, T. L., Zhu, J. Q., Grant, E., & Thomas McCoy, R. (2024). Bayes in the age of intelligent machines. *Current Directions in Psychological Science, 33(5)*, 283-291

  Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science, 331*(6022), 1279-1285

Week 6 ————————————

Class 11 (Monday March 3) — Language Processing

[No Computational Lab]

- Word Perception and Verb Learning (Cohen)
- Rules, Generalization, and Language Models (Goldberg)

  * **Neurobiology of Language.** Hagoort, P. (2019). The neurobiology of language beyond single-word processing. Science, 366(6461), 55-58.

  * **Past Tense Model**. Rumelhart, D. E. & McClelland, J. L. (1986) On Learning the Past Tenses of English Verbs. PDP, Chapter 18. Rumelhart, D. E., & McClelland, J. L..

  Rules and Exceptions. Pinker, S. & Prince, A. (1988) On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28, 73-193.

  * **Letters and Words.** McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological review*, *88*(5), 375.

  PDP Model of Word Reading. Seidenberg, M. S. & McClelland, J. L. (1989) A distributed developmental model of word recognition and naming. *Psychological Review, 96*, 523-568.

Class 12 (Thursday March 6) — Brain Systems, Semantics and Context Inference

- Discourse Processing, Context Inference and Semantics (Cohen)

  * **Deep Learning and Natural Language Processing**. Hirschberg, J., & Manning, C. D. (2015). Advances in natural language processing. *Science*, *349*(6245), 261-266.

  Syntax in LLMs. Hewitt, J., & Manning, C. D. (2019, June). A structural probe for finding syntax in word representations. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1* (Long and Short Papers) (pp. 4129-4138).

  Latent Semantic Analysis. Dumais, S. T. (2004). Latent semantic analysis. Annual Review of Information Science and Technology (ARIST), 38, 189-230.

Word2Vec: *Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. Advances in Neural Information Processing Systems, 26.*

BERT. Tenney, I, Das D & Pavlick E (2019). BERT rediscovers the classical NLP pipeline. *arXiv preprint arXiv:1905.05950*.

Spring Break

Week 7

SECTION 5: EPISODIC MEMORY - BINDING

[LAB: EGO Model]

Class 13 (Monday March 17) — **Episodic Memory and Hippocampal Function**

- Complementary Learning Systems and Hippocampal Function (Cohen)
- Non-Monotonic Plasticity, Memory Sharpening & Sleep (Norman)

* **Episodic memory.** Clewett D., DuBrow S., & Davachi L. (2019). Transcending time in the brain: How event memories are constructed from experience. *Hippocampus*, 29:162-183.

Episodic memory. Tulving, E. (2002). Episodic memory: From mind to brain. *Annual review of Psychology, 53*(1), 1-25.

* **Complementary learning systems.** McClelland J.L., McNaughton B.L., & O'Reilly R.C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. The Psychological Review, 102:419-457.

Complementary learning systems (extended). O'Reilly R.C., & Norman K.A. (2002). Hippocampal and neocrotical contributions to memory: Advances in the complementary learning systems framework. Trends in Cognitive Sciences, 6(12):505-510.

Temporal Context Model. Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. Journal of mathematical psychology, 46(3), 269-299.

CLS theory updated. Kumaran, D., Hassabis, D., & McClelland, J. L. (2016). What learning systems do intelligent agents need? Complementary learning systems theory updated. Trends in cognitive sciences, 20(7), 512-534.

Latent cause inference. Gershman, S. J., & Niv, Y. (2010). Learning latent structure: carving nature at its joints. Current opinion in neurobiology, 20(2), 251-256.

Oscillatory inhibition. *Norman, K. A., Newman, E. L., & Detre, G. (2007). A neural network model of retrieval-induced forgetting. Psychological review, 114(4), 887.*

BCM Learning Rule. Bienenstock, E. L., Cooper, L. N., & Munro, P. W. (1982). Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual *cortex. Journal of Neuroscience, 2*(1), 32-48.

SECTION 6: ATTENTION, WORKING MEMORY AND COGNITIVE CONTROL - MODULATION

Class 14 (Thursday March 20) — Attention, Control and Prefrontal Function

- Attention, Automaticity and Control (Cohen)

* **Empirical Study of Automaticity, Attention and Control.** Shiffrin, RM & Schneider W (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological review*, *84*(2), 127.

* **Prefrontal Cortex and Control.** Miller EK & Cohen JD (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24,* 167-202.

Forms of Attention. Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology, 62(*1), 73-101.

Connectionist Model of Attention, Automaticity and Control. Cohen JD, Dunbar K & McClelland JL (1990). On the control of automatic processes: A parallel distributed processing model of the Stroop effect. *Psychological Review, 97*(3), 332-361.

Week 8 ———————————

[LAB: Stroop Model]

Class 15 (Monday March 24) — Working memory and Prefrontal Function

- Active maintenance (Buschman)
- The dynamics of control (Buschman)

TODO: Bouchacourt et al. (is the paper below what you had in mind?)

Bouchacourt, F., Tafazoli, S., Mattar, M. G., Buschman, T. J., & Daw, N. D. (2022). Fast rule switching and slow rule updating in a perceptual categorization task. Elife, 11, e82531.

TODO: Fusi

Panichello, M. F., DePasquale, B., Pillow, J. W., & Buschman, T. J. (2019). Error-correcting dynamics in visual working memory. Nature communications, 10(1), 3366.

Class 16 (Thursday March 27) — Representation and Capacity Limits

- Compositional vs. Conjunctive coding, Multitasking and Miller's Law (Cohen)

* ***Normative Accounts of Capacity Constraints Associated with Cognitive Control.*** *Musslick S & Cohen JD (2021). Rationalizing constraints on the capacity for cognitive control. Trends in Cognitive Sciences, 25(9), 757-775. https://doi.org/10.1016/j.tics.2021.06.001.*

* **Attention and Binding.** Treisman, A. (1998). Feature binding, attention and object perception. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 353(1373), 1295-1306.

Shared vs. Separated Representation and Serial vs. Parallel Processing. Musslick S, Saxe AM, Hoskin AN, Sagiv Y, Reichman D, Petri G & Cohen JD (under review)**.** On the rational boundedness of cognitive control: Shared versus separated representations. https://psyarxiv.com/jkhdf.

Principled Account of Capacity Constraints in Terms of Representation. Frankland SM, Webb TW, Lewis RL & Cohen JD (in preparation). No coincidence, George: Processing limits in cognitive function reflect the curse of generalization. https://doi.org/10.31234/osf.io/cjuxb.

Week 9 ———————————

Class 17 (Monday March 31) — **Higher Level Cognition**

[LAB: ISC-CI Model]

• **Integration, Attention, Context and Control** (Cohen)

• **Abstraction** (Cohen)

Episodic vs. Working Memory. Beukers AO, Hamin M, Norman KA & Cohen JD (2024). When working memory may be just working, not memory. *Psychological Review, 131*(2), 563. https://psyarxiv.com/jtw5p

* **Semantics, Context and Control. Giallanza T, Campbell D, Cohen JD & Rogers TT (2024).** An integrated model of semantics and control. *Psychological Review.* https://psyarxiv.com/jq7ta.

Semantics, Control and Context Inference. Giallanza T, Rogers TT & Cohen JD. (under review). An integrated model of semantics and control, Part 2: Solving the similarity paradox through context inference. https://osf.io/preprints/psyarxiv/fxc87.

External memory in neural networks: The Neural Turing Machine. Graves, A. (2014). Neural Turing Machines. arXiv preprint arXiv:1410.5401.

Episodic Memory and Abstraction. Webb TW, Sinha I & Cohen JD (2021). Emergent symbols through binding in external memory. *ICLR 2021: Proceedings of the International Conference on Learning Representations*. https://arxiv.org/abs/2012.14601

Relational Bottleneck. Webb TW, Frankland SM, Altabaa A, Segert S, Krishnamurthy K, Campbell D, Russin J, Giallanza T, Dulberg Z, Reilly RO, Lafferty J & Cohen JD (2024). The Relational bottleneck as an inductive bias for efficient abstraction. *Trends in Cognitive Science, 28(9)*:829-843. doi: 10.1016/j.tics.2024.04.001; https://arxiv.org/abs/2309.06629.

* **Episodic Generalization and Control.** Giallanza T, Campbell D & Cohen JD (2024). Toward the emergence of intelligent control: Episodic generalization and optimization. *OpenMind, 8,* 688-722. https://osf.io/preprints/psyarxiv/dzvpy/

* **Connectionist vs. Symbolic Processing.** Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*(1-2), 3-71.

Connectionist Model of Symbol Processing. Touretzky, D. S., & Hinton, G. E. (1985). Symbols among the neurons: Details of a connectionist inference architecture. In *IJCAI* (Vol. 85, pp. 238-243).

Analogical Reasoning and LLMs. Webb, T., Holyoak, K.J. & Lu, H. Emergent analogical reasoning in large language models. *Nat Hum Behav* **7**, 1526–1541 (2023). https://doi.org/10.1038/s41562-023-01659-w

In-context Learning and Induction Heads. Olsson, C., Elhage, N., Nanda, N., Joseph, N., DasSarma, N., Henighan, T., ... & Olah, C. (2022). In-context learning and induction heads. arXiv preprint arXiv:2209.11895.

Tensor Product Representations. Plate, T. A. (1995). Holographic reduced representations. *IEEE Transactions on* Neural networks, 6(3), 623-641.

Tensor Product Representations and Symbolic Processing. Smolensky, P. (1990). Tensor product variable binding and the representation of symbolic structures in connectionist systems. *Artificial intelligence*, *46*(1-2), 159-216.

Tensor product representations and Deep Learning. McCoy, R. T., Linzen, T., Dunbar, E., & Smolensky, P. (2018). RNNs implicitly implement tensor product representations. arXiv preprint arXiv:1812.08718.

LLM content effects. Dasgupta, I., Lampinen, A. K., Chan, S. C., Creswell, A., Kumaran, D., McClelland, J. L., & Hill, F. (2022). Language models show human-like content effects on reasoning. arXiv preprint arXiv:2207.07051, 2(3).

Class 18 (Thursday April 3) — **Optimization and Control**

• Performance Monitoring and the Expected Value of Control (Cohen)

* **Conflict Monitoring.** Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. Psychological review, 108(3), 624.

* **Expected Value of Control.** Shenhav A, Botvinick MM & Cohen JD (2013). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron, 79*, 217-240.

* **Resource rational analysis**. Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. Behavioral and brain sciences, 43, e1.

Bounded Rationality. Howes, A., Lewis, R. L., & Vera, A. (2009). Rational adaptation under task and processing constraints: implications for testing theories of cognition and action. *Psychological review*, *116*(4), 717.

Resource Rationality and Effort. Shenhav A, Musslick S, Lieder F, Kool W, Griffiths TL, Cohen JD & Botvinick MM (2017). Toward a rational and mechanistic account of mental effort. *Annual Review of Neuroscience, 40*, 99-124.

Week 10 ————————————

SECTION 7: MOTOR FUNCTION

[LAB: Sequential adjustment effects / EVC model]

Class 19 (Monday April 7) — **Motor Representations and Planning**

(Graziano)

(Buschman)

Class 20 (Thursday April 10) **— Motor Control, Learning and Cerebellar Function**

(Taylor)

Week 11 ————————————

SECTION 8: DEVELOPMENT and SOCIAL COGNITION

Class 21 (Monday April 14) — **Development**

Infantile amnesia. Howe, M. L., & Courage, M. L. (1993). On resolving the enigma of infantile amnesia. Psychological bulletin, 113(2), 305.

Baillargeon, R. (1987). Object permanence in 3½-and 4½-month-old infants. Developmental psychology, 23(5), 655.

Infantile Amnesia. Ellis, C. T., Skalaban, L. J., Yates, T. S., Bejjanki, V. R., Córdova, N. I., & Turk-Browne, N. B. (2021). Evidence of hippocampal learning in human infants. Current Biology, 31(15), 3358-3364.

Class 22 (Thursday April 17) **— Social Cognition**

(Hasson)

Neural coupling. Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., & Keysers, C. (2012). Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trends in cognitive sciences*, *16*(2), 114-121.

AK-to-JDC: Please verify this is what you meant by "Dynamic programming models"

Bayesian model of pragmatics. Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. Trends in cognitive sciences, 20(11), 818-829.

AK-to-JDC: Please verify this is what you meant by "Theory of Mind - Baron-Cohen"

Theory of mind and autism. Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"?. Cognition, 21(1), 37-46

Brain regions responsible for social cognition. Saxe, R. (2006). Uniquely human social cognition. Current opinion in neurobiology, 16(2), 235-239.

Week 12

————————————

SECTION 9: DISORDERS

Class 23 (Monday April 21) — Cognitive Neuropsychology **(Kastner)** / Origins of Computational Psychiatry

Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: a connectionist approach to behavior and biology in schizophrenia. Psychological review, 99(1), 45.

Class 24 (Thursday April 24) — Modern Computational Psychiatry (Cohen/**Niv**)

Pisupati, S., Langdon, A. J., Konova, A. B., & Niv, Y. (2024). The utility of a latent-cause framework for understanding addiction phenomena. *Addiction Neuroscience,* 10, 100143.

AK-to-JDC: Please verify this is what you meant by "Hartley & Daw - big data"

Decker, J. H., Otto, A. R., Daw, N. D., & Hartley, C. A. (2016). From creatures of habit to goal-directed learners: Tracking the developmental emergence of model-based reinforcement learning. Psychological science, 27(6), 848-858.