Curriculum Vitae ***Samuel J. Gershman***

PERSONAL DETAILS

Samuel J. Gershman
Department of Psychology and Center for Brain Science
Harvard University
Northwest Building, room 190.06
52 Oxford St.
Cambridge, MA 02138
gershman@fas.harvard.edu
773-607-9817

EMPLOYMENT HISTORY

Professor	2021-
Associate Professor	2019-2021
Assistant Professor	2015-2019
Department of Psychology and Center for Brain Science, Harvard University	
Postdoctoral fellow	2013-2015
Department of Brain and Cognitive Sciences, MIT	
(advisors: Joshua Tenenbaum & Nancy Kanwisher)	
Research assistant, NYU	2007-2008

EDUCATION

BA	Neuroscience and Behavior Columbia University	2003-2007
PhD	Psychology and Neuroscience Princeton University (advisors: Kenneth Norman & Yael Niv)	2008-2013

FELLOWSHIPS and AWARDS

2024	Schmidt Sciences Polymath Award
2023	Election to the Society of Experimental Psychology
2021	Walter Channing Cabot Fellowship
2020	Janet Taylor Spence Award for Transformative Early Career Contributions from the American Psychological Society
2020	Cognitive Neuroscience Society Young Investigator Award
2018	Bjorkman-Strominger-Wiley prize for collaboration from the Harvard Molecular and Cellular Biology Department (joint with Nao Uchida)
2018	Alfred P. Sloan Research Fellowship
2015	Clever Systems Early Career Investigator Award from the APA
2014	Glushko Dissertation Award from the Cognitive Science Society

2013 MIT Intelligence Initiative Postdoctoral Fellowship
2012 International Conference for Machine Learning travel award
2010 NSF graduate research fellowship
2009 NIPS travel award
2009 NIMH Quantitative and Computational Neuroscience training fellowship
2009 Swartz COSYNE travel fellowship
2008 Walker McKinney ’50 Life Sciences Fellowship
2006 Summer University Research Fellowship, Columbia University

PUBLICATIONS

1. Irie, K., Yau, M., & Gershman, S.J. (in press). Blending complementary memory systems in hybrid quadratic-linear transformers. *Advances in Neural Information Processing Systems*, 38.
2. Martinez, J.E., Krasner, R.H., Rosero, L., Gershman, S.J., & Cikara, M. (in press). Social group discovery, structure, and stereotype updating. *Journal of Experimental Psychology: General.*
3. Fan, H., Callaway, F., & Gershman, S.J. (in press). Uncertainty-driven exploration during planning. *Decision.*
4. Xiang, Y., Gershman, S.J., & Gerstenberg, T. (in press). A signaling theory of self-handicapping. *Cognition*.
5. Hamou, N., Gershman, S.J., & Reddy, G. (in press). Reconciling time and prediction error theories of associative learning. *Nature Communications.*
6. Tsividis, P.A., Loula, J., Burga, J., Foss, N., Campero, A., Pouncy, T., Gershman, S.J., & Tenenbaum, J.B. (in press). Human-level reinforcement learning through theory-based modeling, exploration, and planning. *Philosophical Transactions A.*
7. Battleday, R.M., & Gershman, S.J. (in press). Artificial intelligence for science: the easy and hard problems. *Philosophical Transactions A.*
8. Bari, B.A., Krystal, A.D., Pizzagalli, D.A., & Gershman, S.J. (in press). Computationally-informed insights into anhedonia and treatment by κ-opioid receptor antagonism. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*.
9. Liu, S., Kikumoto, A., Badre, D., & Gershman, S.J. (2025). Neural and behavioral signatures of policy compression in cognitive control. *Cerebral Cortex, 35,* bhaf223.
10. Masset, P., & Gershman, S.J. (2025). Reinforcement learning with dopamine: a convergence of natural and artificial intelligence. In S. Cragg & M. Walton (eds). *The Handbook of Dopamine*. Elsevier.
11. Gershman, S.J., Bill, J., & Drugowitsch, J. (2025). Hierarchical vector analysis of visual motion perception. *Annual Review of Vision Science, 11,* 411-422.
12. Prat-Carrabin, A., & Gershman, S.J. (2025). Bayesian estimation yields anti-Weber variability. *PNAS Nexus, 4*, pgaf275.

13. Gershman, S.J. (2025). Bridging computation and representation in associative learning. *Computational Brain and Behavior, 8,* 377-391.
14. Liu, S., & Gershman, S.J. (2025). Action subsampling supports policy compression in large action spaces. *PLOS Computational Biology, 21*, e1013444.
15. Lai, L., Huang, A.Z., & Gershman, S.J. (2025). Action chunking as conditional policy compression. *Cognition, 264,* 106201.
16. Y., Dorst, K., & Gershman, S.J. (205). On the robustness and provenance of the gambler's fallacy. *Psychological Science, 36,* 451-464.
17. Qu, A.J., Tai, L-H., Hall, C.D., Tu, E.M., Eckstein, M.K., Mischanchuk, K., Lin, W.C., Chase, J.B., MacAskill, A.F., Collins, A.G.E., Gershman, S.J., & Wilbrecht, L. (2025). Nucleus accumbens dopamine release reflects Bayesian inference during instrumental learning. *PLOS Computational Biology, 21*, e1013226.
18. Ahmed, Z., Tenenbaum, J.B., Bates, C.J., & Gershman, S.J. (2025). Synthesizing world models for bilevel planning. *Transactions on Machine Learning Research.*
19. Qian, L., Burrell, M., Hennig, J.A., Matias, S., Murthy, V.N., Gershman, S.J., & Uchida, N. (2025). The role of prospective contingency in the control of behavior and dopamine signals during associative learning. *Nature Neuroscience, 28,* 1280-1292.
20. Jaroszewski, A.C., Millner, A.J., Gershman, S.J., Franz, P.J., Bentley, K.H., Kleiman, E.M., & Nock, M.K. (2025). Past suicide attempt is associated with a weaker decision-making bias to actively escape from suicide-related stimuli. *Journal of Psychopathology and Clinical Science, 28,* 1280-1292.
21. Liu, S., Lai, L., Gershman, S.J., & Bari, B.A. (2025). Time and memory costs jointly determine a speed-accuracy trade-off and set-size effects. *Journal of Experimental Psychology: General, 154,* 1611-1627.
22. Gershman, S.J., Fiete, I., & Irie, K. (2025). Key-value memory in the brain. *Neuron, 113,* 1694-1707.
23. Hall-McMaster, S., Tomov, M., Gershman, S.J., & Schuck, N.W. (2025). Neural evidence that humans reuse strategies to solve new tasks. *PLOS Biology, 23*, e3003174.
24. Gershman, S.J., & Lak, A. (2025). Policy complexity suppresses dopamine responses. *Journal of Neuroscience, 45, e1756242024*.
25. Han, S., Pari, J., Gershman, S.J., & Agrawal, P. (2025). General reasoning requires learning to reason from the get-go. *Proceedings of the 42nd International Conference on Machine Learning*.
26. Liu, S., Gershman, S.J., & Bari, B. (2025). Quantifying the cost of context sensitivity in decision making. *Proceedings of the 47th Annual Cognitive Science Society.*

27. Ying, L., Truong, R., Tenenbaum, J.B., & Gershman, S.J. (2025). Adaptive social learning using theory of mind. *Proceedings of the 47th Annual Cognitive Science Society.*

28. Xiang, Y., & Gershman, S.J. (2025). Modeling intrinsic motivation as reflective planning. *Proceedings of the 47th Annual Cognitive Science Society.*

29. Xiang, Y., Bigelow, E., Gerstenberg, T., Ullman, T., & Gershman, S.J. (2025). Language models assign responsibility based on actual rather than counterfactual contributions. *Proceedings of the 47th Annual Cognitive Science Society.*

30. Bari, B.A., & Gershman, S.J. (2025). The value of non-instrumental information in anxiety: insights from a resource-rational model of planning. *Computational Psychiatry, 9*, 63-75.

31. Fry, B.R., Russell, N., Fex, V., Mo, B., Pence, N., Beatty, J.A., Manfreddsson, F.P., Toth, B.A., Burgess, C.R., Gershman, S.J., & Johnson, A.W. (2025). Devaluing memories of reward: A case for dopamine. *Communications Biology, 8*, 161.

32. Binz et al. (2025). How should the advent of large language models affect the practice of science? *Proceedings of the National Academy of Science, 122*, e2401227121.

33. Kumar, T., Bordelon, B., Pehlevan, C., Murthy, V.N., & Gershman, S.J. (2025). Do mice grok? glimpses of hidden progress during overtraining in sensory cortex. *International Conference on Learning Representations 13*.

34. Xiang, Y., Landy, J., Cushman, F.A., Velez, N., & Gershman, S.J. (2025). People reward others based on their willingness to exert effort. *Journal of Experimental Social Psychology, 116*, 104699.

35. Sosa, F., Gershman, S.J., & Ullman, T. (2025). Blending simulation and abstraction for physical reasoning. *Cognition, 254*, 105995.

36. Nguyen, T., Bezdek, M., Gershman, S.J., Bobick, A.F., Braver, T., & Zacks, J.M. (2024). Modeling human activity comprehension at human scale: Prediction, segmentation, and categorization. *PNAS Nexus, 3*, 459.

37. Carvalho, W., Tomov, M.S., de Cothi, W., Barry, C., & Gershman, S.J. (2024). Predictive representations: building blocks of intelligence. *Neural Computation, 36*, 2225-2298.

38. Coppersmith, D.D.L., Jaroszewski, A.C., Gershman, S.J., Cha, C.B., Millner, A.J., Fortgang, R.G., Kleiman, E.M., & Nock, M.K. (2024). Do people know how suicidal they will be? Understanding suicidal prospection. *Suicide and Life-Threatening Behavior, 54,* 750-761.

39. Lu, Q., Nguyen, T.T., Zhang, Q., Hasson, U., Griffiths, T.L., Zacks, J.M., Gershman, S.J., & Norman, K.A. (2024). Reconciling shared versus context-specific information in a neural network model of latent causes. *Scientific Reports, 14*, 16782.

40. Orchinik, R., Dubey, R., Gershman, S.J., Powell, D., Bhui, R. (2024). Learning from and about scientists: Consensus messaging shapes perceptions of climate change and climate scientists. *PNAS Nexus, 3*, 485.
41. Chen, A.M., Palacci, A., Velez, N., Hawkins, R., & Gershman, S.J. (2024). A hierarchical Bayesian approach to adaptive teaching. *Cognitive Science, 48,* e13477.
42. Gershman, S.J. (2024). Habituation as optimal filtering. *iScience, 27*, 110523.
43. Gershman, S.J., Assad, J.A., Datta, S.R., Linderman, S.W., Sabatini, B.L., Uchida, N., & Wilbrecht, L. (2024). Explaining dopamine through prediction errors and beyond. *Nature Neuroscience, 27*, 1645-1655.
44. Lai, L., & Gershman, S.J. (2024). Human decision making balances reward maximization and policy compression. *PLOS Computational Biology*, *20*, e1012057.
45. Bari, B.A., & Gershman, S.J. (204). Resource-rational psychopathology. *Behavioral Neuroscience*, *138*, 221-234.
46. Gershman, S.J. (2024). What have we learned about artificial intelligence from studying the brain? *Biological Cybernetics, 118,* 1-5.
47. Schurr, R., Reznik, D., Hillman, H., Bhui, R., & Gershman, S.J. (2024). Dynamic computational phenotyping of human cognition. *Nature Human Behaviour, 8,* 917-931.
48. Allen, K., Brändle, F., ... Schulz, E. (in press) Using games to understand the mind. *Nature Human Behaviour*.
49. Zhao, J., Radke, J., Chen, F.S., Sachdeva, S., Gershman, S.J., & Luo, Y. (2024). How do we reinforce climate action? *Sustainability Science, 19,* 1503-1517.
50. Geerts, J.P., Gershman, S.J., Burgess, N., & Stachenfeld, K.L. (2024). A probabilistic successor representation for context-dependent prediction. *Psychological Review, 131,* 578-597.
51. Beukers, A.O., Collin, S.H.P., Kempner, R.P., Franklin, N.T., Gershman, S.J., & Norman, K.A. (2024). Blocked training facilitates learning of multiple schemas. *Communications Psychology, 2*, 28.
52. Xiang, Y., Velez, N., & Gershman, S.J. (2024). Optimizing competence in the service of collaboration. *Cognitive Psychology, 150*, 101653.
53. Kumar, T., Bordelon, B., Gershman, S.J., & Pehlevan, C. (2024). Grokking as the transition from lazy to rich training dynamics. *International Conference on Learning Representations 12*.
54. Bates, C.J., Alvarez, G., & Gershman, S.J. (2024). Scaling models of visual working memory to natural images. *Communications Psychology, 2*, 3.
55. Hennig, J., Romeiro Pinto, S.A., Yamaguchi, T., Linderman, S.W., Uchida, N., & Gershman, S.J. (2023). Emergence of belief-like representations through reinforcement learning. *PLOS Computational Biology, 19,* e1011067.

56. Li, Y., Wang, Y., Boger, T., Smith, K., Gershman, S.J., & Ullman, T.D. (2023). An approximate representation of objects underlies physical reasoning. *Journal of Experimental Psychology: General, 152,* 3074-3086.

57. Brändle, F., Stocks, L.J., Tenenbaum, J.B., Gershman, S.J., & Schulz, E. (2023). Intrinsic exploration as empowerment in a richly structured online game. *Nature Human Behaviour, 7,* 1481-1489.

58. Gershman, S.J., & Cikara, M. (2023). Structure learning principles of stereotype change. *Psychonomic Bulletin & Review, 30,* 1273-1293.

59. Xiang, Y., Velez, N., & Gershman, S.J. (2023). Collaborative decision making is grounded in representations of other people's competence and effort. *Journal of Experimental Psychology: General, 152,* 1565-1579.

60. Xiang, Y., Landy, J., Cushman, F.A., Velez, N., & Gershman, S.J. (2023). Actual and counterfactual effort contribute to responsibility attributions in collaborative tasks. *Cognition, 241*, 105609.

61. Velez, N., Chen, A.M., Burke, T., Cushman, F.A., & Gershman, S.J. (2023). Teachers recruit mentalizing regions to represent learners' beliefs. *Proceedings of the National Academy of Sciences, 120*, e2215015120.

62. Gershman, S.J., & Burke, T. (2023). Mental control of uncertainty. *Cognitive, Affective, and Behavioral Neuroscience, 23,* 465-475.

63. Jakob, A., & Gershman, S.J. (2023). Rate-distortion theory of neural coding and its implications for working memory. *eLife, 12*, e79450.

64. Fan, H., Gershman, S.J., & Phelps, E.A. (2023). Trait somatic anxiety is associated with reduced directed exploration and underestimation of uncertainty. *Nature Human Behaviour, 7,* 102-113.

65. Fan, H., Burke, T., Sambrano, D., Dial, E., Phelps, E.A., & Gershman, S.J. (2023). Pupil size encodes uncertainty during exploration. *Journal of Cognitive Neuroscience, 35, 1508-1520*.

66. Bari, B.A., & Gershman, S.J. (2023). Undermatching is a consequence of policy compression. *Journal of Neuroscience, 43,* 447-457.

67. Gershman, S.J., & Ullman, T.D. (2023). Causal implicatures from correlational statements. *PLOS One, 18*, e0286067.

68. Gershman, S.J. (2023). The rational analysis of memory. In M. Kahana & A. Wagner, Eds, *Oxford Handbook of Human Memory*. Oxford University Press.

69. Tomov, M.S., Tsividis, P.A., Pouncy, T., Tenenbaum, J.B., & Gershman, S.J. (2023). The neural architecture of theory-based reinforcement learning. *Neuron, 111,* 1331-1344.

70. Biderman, N., Gershman, S.J., Shohamy, D. (2023). The role of memory in counterfactual valuation. *Journal of Experimental Psychology: General, 152*, 1754-1767.

71. Gershman, S.J. (2023). The molecular memory code and synaptic plasticity: a synthesis. *BioSystems*, *224*, 104825.
72. Orchinik, R., Dubey, R., Gershman, S.J., Powell, D., Bhui, R. (2023). Learning about scientists from climate consensus messaging. *Proceedings of the 45th Annual Cognitive Science Society.*
73. Binz, M., Gershman, S.J., Schulz, E., & Endres, D. (2022). Heuristics from bounded meta-learned inference. *Psychological Review, 129,* 1042-1077.
74. McNamee, D., Stachenfeld, K.L., Botvinick, M.M., Gershman, S.J. (2022). Compositional sequence generation in the entorhinal-hippocampal system. *Entropy, 24*, 1791.
75. Pouncy, T., & Gershman, S.J. (2022). Inductive biases in theory-based reinforcement learning. *Cognitive Psychology, 138*, 101509.
76. Alexander, W.H., & Gershman, S.J. (2022). Representation learning with reward prediction errors. *Neurons, Behavior, Data Analysis, and Theory*.
77. Bill, J., Gershman, S.J., & Drugowitsch, J. (2022). Visual motion perception as online hierarchical inference. *Nature Communications, 13*, 7403.
78. Mikhael, J.G., & Gershman, S.J. (2022). Impulsivity and risk-seeking as Bayesian inference under dopaminergic control. *Neuropsychopharmacology, 47,* 465-476.
79. Bates, C.J., & Gershman, S.J. (2022). Coding strategies in memory for 3D objects: the influence of task uncertainty. *Proceedings of the 44th Annual Conference of the Cognitive Science Society*, 1077-1087.
80. Mikhael, J.G., Kim, H.R., Uchida, N., & Gershman, S.J. (2022). The role of state uncertainty in the dynamics of dopamine. *Current Biology, 32*, 1077-1087.
81. Le, T.A., Collins, K.M., Hewitt, L., Ellis, L., Siddharth, N., Gershman, S.J., & Tenenbaum, J.B. (2022). Hybrid memoised wake-sleep: approximate inference at the discrete-continuous interface. *10th International Conference on Learning Representations*.
82. Wu, C.M., Schulz, E., & Gershman, S.J. (2021). Inference and search on graph-structured spaces. *Computational Brain and Behavior, 4,* 125-147.
83. Gershman, S.J. (2021). Just looking: the innocent eye in neuroscience. *Neuron*, *109,* 2220-2223.
84. Gershman, S.J. (2021). *What Makes Us Smart: The Computational Logic of Human Cognition.* Princeton University Press. Princeton: NJ.
85. McNamee, D., Stachenfeld, K.L., Botvinick, M.M., Gershman, S.J. (2021). Flexible modulation of sequence generation in the entorhinal–hippocampal system. *Nature Neuroscience*, *24*, 851-862.
86. Sosa, F.A., Ullman, T., Tenenbaum, J.B., Gershman, S.J., & Gerstenberg, T. (2021). Moral dynamics: grounding moral judgment in intuitive theories. *Cognition, 217*, 104890.
87. Bhattasali, N.X., Tomov, M.S., & Gershman, S.J. (2021). CCNLab: A benchmarking framework for computational cognitive neuroscience. *35th Conference on Neural Information Processing Systems (NeurIPS 2021) Track on Datasets and Benchmarks.*

88. Dorfman, H.M., Tomov, M., Cheung, B., Clarke, D., Gershman, S.J., & Hughes, B.L. (2021). Causal inference gates corticostriatal learning. *Journal of Neuroscience, 41*, 6892-6904.
89. Lai, L., & Gershman, S.J. (2021). Policy compression: an information bottleneck in action selection. *Psychology of Learning and Motivation*, *74*, 195-232.
90. Tomov, M., Schulz, E., & Gershman, S.J. (2021). Multi-task reinforcement learning in humans. *Nature Human Behaviour, 5*, 764-773.
91. Xiang, Y., Graeber, T., Enke, B., & Gershman, S.J. (2021). Confidence and central tendency in perceptual judgment. *Attention, Perception, & Psychophysics, 83*. 3024-3034.
92. Gershman, S.J., Guitart-Masip, M., & Cavanagh, J.F. (2021). Neural signatures of arbitration between Pavlovian and instrumental action selection. *PLOS Computational Biology, 17,* e1008553.
93. Gershman, S.J., & Lai, L. (2021). The reward-complexity trade-off in schizophrenia. *Computational Psychiatry, 5,* 38-53.
94. Dasgupta, I., & Gershman, S.J. (2021). Memory as a computational resource. *Trends in Cognitive Sciences, 25*, 240-251.
95. Bhui, R., Lai, L., & Gershman, S.J. (2021). Resource-rational decision making. *Current Opinion in Behavioral Sciences, 41,* 15-21.
96. Mikhael, J.G., Lai, L., & Gershman, S.J. (2021). Rational inattention and tonic dopamine. *PLOS Computational Biology, 17*, e1008659.
97. Yang, S., Bill, J., Drugowitsch, J., & Gershman, S.J. (2021). Human visual motion perception shows hallmarks of Bayesian structural inference. *Scientific Reports, 11, 3714*.
98. Gershman, S.J., Balbi, P.E.M., Gallistel, C.R., & Gunawardena, J. (2021). Reconsidering the evidence for learning in single cells. *eLife, 10*, e61907.
99. Pouncy, T., Tsividis, P., & Gershman, S.J. (2021). What is the model in model-based planning? *Cognitive Science, 45,* e12928.
100. Gershman, S.J. (2020). Origin of perseveration in the trade-off between reward and complexity. *Cognition, 204, 104394*.
101. Dasgupta, I., Guo, D., Gershman, S.J., Goodman, N.D. (2020). Analyzing machine-learned representations: A natural language case study. *Cognitive Science, e12925*.
102. Bill, J. Pailian, H., Gershman, S.J., & Drugowitsch, J. (2020). Hierarchical structure is employed by humans during visual motion perception. *Proceedings of the National Academy of Sciences, 117,* 24581-24589.
103. Kim, H.R., Malik, A.N., Mikhael, J.G., Bech, P., Tsutsui-Kimura, I., Sun, F., Zhang, Y., Li, Y., Watabe-Uchida, M., Gershman, S.J., & Uchida, N. (2020). A unified framework for dopamine signals across timescales. *Cell, 183,* 1600-1616.
104. Cohen, A.O., Nussenbaum, K., Dorfman, H.M., Gershman, S.J., & Hartley, C.A. (2020). The rational use of causal inference to guide reinforcement learning changes with age. *NPJ Science of Learning, 5,* 16.
105. Bhui, R., & Gershman, S.J. (2020). Paradoxical effects of persuasive messages. *Decision, 7,* 239-258.
106. Gershman, S.J. & Cikara, M. (2020). Social structure learning. *Current Directions in Psychological Science*, *29*, 460-466.

107. Goldwater, M.B., Gershman, S.J., Moul, C., Ludowici, C., Burton, A., Killer, B., Kuhnert, R-L., Ridgway, K. (2020). Children's understanding of habitual behaviour. *Developmental Science, 23,* e12951.
108. Schulz, E., Quiroga, F., & Gershman, S.J. (2020). Communicating compositional patterns. *Open Mind, 4,* 25-39.
109. Baumann, C., Singmann, H., Gershman, S.J., & von Helversen, B. (2020). A linear threshold model for optimal stopping behavior. *Proceedings of the National Academy of Sciences, 117*, 12750-12755.
110. Tomov, M., Truong, V., Hundia, R., & Gershman, S.J. (2020). Dissociable neural correlates of uncertainty underlie different exploration strategies. *Nature Communications, 11,* 2371.
111. Gershman, S.J., & Olveczky, B.P. (2020). The neurobiology of deep reinforcement learning. *Current Biology, 30*, R617-R634.
112. Sanders, H., Wilson, M.A., Gershman, S.J. (2020). Hippocampal remapping as hidden state inference. *eLife, 9*, e51140.
113. Lau, T., Gershman, S.J., & Cikara, M. (2020). Social structure learning in human anterior insula. *eLife,* e53162.
114. Gershman, S.J. & Bhui, R. (2020). Rationally inattentive intertemporal choice. *Nature Communications, 11*, 3365.
115. Tomov, M., Yagati, S., Kumar, A., Yang, W., & Gershman, S.J. (2020). Discovery of hierarchical representations for efficient planning. *PLOS Computational Biology*, *16*, e1007594.
116. Dasgupta, I., Schulz, E., Tenenbaum, J.B. & Gershman, S.J. (2020). A theory of learning to infer. *Psychological Review*, *127*, 412-441.
117. Franklin, N.T., Norman, K.A., Ranganath, C., Zacks, J.M., & Gershman, S.J. (2020). Structured event memory: a neuro-symbolic model of event cognition. *Psychological Review , 127,* 327-361.
118. Schulz, E., Franklin, N.T. & Gershman, S.J. (2020). Finding structure in multi-armed bandits. *Cognitive Psychology, 119,* 101261.
119. Dorfman, H.M., & Gershman, S.J. (2019). Controllability governs the balance between Pavlovian and instrumental action selection. *Nature Communications*, *10*, 5826.
120. Stalnaker, T., Howard, J., Takahashi, Y., Gershman, S.J., Kahnt, T., & Schoenbaum, G. (2019). Dopamine neuron ensembles signal the content of sensory prediction errors. *eLife*, e49315.
121. Gershman, S.J. (2019). What does the free energy principle tell us about the brain? *Neurons, Behavior, Data Analysis, and Theory*.
122. Gershman, S.J. & Uchida, N. (2019). Believing in dopamine. *Nature Reviews Neuroscience, 20,* 703-714.
123. Gershman, S.J. (2019). The generative adversarial brain. *Frontiers in Artificial Intelligence, 2,* 18.
124. Schulz, E., Bhui, R., Love, B.C., Brier, B., Todd, M.T., & Gershman, S.J. (2019). Structured, uncertainty-driven exploration in real-world consumer choice. *Proceedings of the National Academy of Sciences, 116,* 13903-13908.
125. Gershman, S.J. (2019). Uncertainty and exploration. *Decision, 6,* 277-286.

126. Mikhael, J.G. & Gershman, S.J. (2019). Adapting the flow of time with dopamine. *Journal of Neurophysiology*, *121*, 1748-1760.
127. Kurdi, B., Gershman, S.J., & Banaji, M.R. (2019). Model-free and model-based learning processes in the updating of explicit and implicit evaluations. *Proceedings of the National Academy of Sciences, 116*, 6035-6044.
128. Cushman, F., & Gershman, S.J. (2019). Editor's introduction: computational approaches to social cognition. *Topics in Cognitive Science, 11*, 281-298.
129. Dorfman, H.M., Bhui, R., Hughes, B.L., & Gershman, S.J. (2019). Causal inference about good and bad outcomes. *Psychological Science, 30,* 516-525.
130. Gershman, S.J. (2019). How to never be wrong. *Psychonomic Bulletin & Review, 26,* 13-28.
131. Tiganj, Z., Gershman, S.J., Sederberg, P.B., & Howard, M.W. (2019). Estimating scale-invariant future in continuous time. *Neural Computation, 31,* 681-709.
132. Patzelt, E., Kool, W., Millner, A.J., & Gershman, S.J. (2019). Incentives boost model-based control across a range of severity on several psychiatric constructs. *Biological Psychiatry, 85,* 425-433.
133. Millner, A.J., den Ouden, H.E.M., Gershman, S.J., Glenn, C.R., Kearns, J., Bornstein, A.M., Marx, B.P., Keane, T.M., & Nock, M.K. (2019). Suicidal thoughts and behaviors are associated with an increased decision-making bias for active responses to escape aversive states. *Journal of Abnormal Psychology, 128,* 106-118.
134. Patzelt, E.H., Kool, W., Millner, A.J., & Gershman, S.J. (2019). The transdiagnostic structure of mental effort avoidance. *Scientific Reports, 9,* 1689.
135. Schulz, E., & Gershman, S.J. (2019). The algorithmic architecture of exploration in the human brain. *Current Opinion in Neurobiology*, *55*, 7-14.
136. Lau, T., Pouncy, H.T., Gershman, S.J., & Cikara, M. (2018). Discovering social groups via latent structure learning. *Journal of Experimental Psychology: General, 147,* 1881-1891.
137. Patzelt, E., Hartley, C.A., & Gershman, S.J. (2018). Computational phenotyping: using models to understand personality, development, and mental illness. *Personality Neuroscience, 1,* e18.
138. Bhui, R., & Gershman, S.J. (2018). Decision by sampling implements efficient coding of psychoeconomic functions. *Psychological Review, 125,* 985-1001.
139. Gardner, M.P.H., Schoenbaum, G., & Gershman, S.J. (2018). Rethinking dopamine as generalized prediction error. *Proceedings of the Royal Society B, 285*, 20181645.
140. Gershman, S.J., & Tzovaras, B.G. (2018). Dopaminergic genes are associated with both directed and random exploration. *Neuropsychologia, 120,* 97-104.
141. Petter, E.A., Gershman, S.J., & Meck, W.H. (2018). Integrating models of interval timing and reinforcement learning. *Trends in Cognitive Sciences*, *22*, 911-922.
142. Millner, A.J., Gershman, S.J., Nock, M.K., & Ouden, H.D. (2018). Pavlovian control of escape and avoidance. *Journal of Cognitive Neuroscience, 30,* 1379-1390.

143. Lage, I., Ross, A.S., Kim, B., Gershman, S.J., & Doshi-Velez, F. (2018). Human-in-the-loop interpretability prior. *Advances in Neural Information Processing Systems 32*.
144. Kool, W., Gershman, S.J., & Cushman, F.A. (2018). Planning complexity registers as a cost in metacontrol. *Journal of Cognitive Neuroscience, 30,* 1391-1404.
145. Tomov, M.S., Dorfman, H.M., & Gershman, S.J. (2018). Neural computations underlying causal structure learning. *Journal of Neuroscience , 38,* 7143-7157.
146. Gershman, S.J. (2018). The successor representation: its computational logic and neural substrates. *Journal of Neuroscience , 38,* 7193-7200.
147. Babayan, B.M., Uchida, N., & Gershman, S.J. (2018). Belief state representation in the dopamine system. *Nature Communications, 9,* 1891.
148. Dasgupta, I., Smith, K.A., Schulz, E., Tenenbaum, J.B., & Gershman, S.J. (2018). Learning to act by integrating mental simulations and physical experiments. *Proceedings of the 40th Annual Conference of the Cognitive Science Society.*
149. Dasgupta, I., Guo, D., Stuhlmuller, A., Gershman, S.J., & Goodman, N.D. (2018). Evaluating compositionality in sentence embeddings. *Proceedings of the 40th Annual Conference of the Cognitive Science Society.*
150. Baumann, C., Singmann, H., Gershman, S.J., & von Helversen, B. (2018). Explaining human decision making in optimal stopping tasks. *Proceedings of the 40th Annual Conference of the Cognitive Science Society.*
151. Gershman, S.J. (2018). Deconstructing the human algorithms for exploration. *Cognition, 173,* 34-42.
152. Starkweather, C.K., Gershman, S.J., & Uchida, N. (2018). Medial prefrontal cortex shapes dopamine reward prediction errors under state uncertainty. *Neuron*, *98,* 616-629.
153. Pereira, F., Lou, B., Pritchett, B., Ritter, S., Gershman, S.J., Kanwisher, N., Botvinick, M., & Fedorenko, E. (2018). Toward a universal decoder of linguistic meaning from brain activation. *Nature Communications, 9,* 963.
154. Kool, W., & Cushman, F.A., & Gershman, S.J. (2018). Competition and cooperation between multiple reinforcement learning systems. In R.W. Morris & A. Bornstein (Eds.) *Goal-Directed Decision Making: Computations and Neural Circuits*. Elsevier.
155. Blanchard, T.C., & Gershman, S.J. (2018). Pure correlates of exploration and exploitation in the human brain. *Cognitive, Affective, and Behavioral Neuroscience, 18,* 117-126.
156. Stachenfeld, K.L., Botvinick, M.M., & Gershman, S.J. (2017). The hippocampus as a predictive map. *Nature Neuroscience, 20,* 1643-1653.
157. Lake, B.M., Ullman, T.D., Tenenbaum, J.B., & Gershman, S.J. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences, 40,* e253.
158. Schulz, E., Tenenbaum, J.B., Duvenaud, D., Speekenbrink, M., & Gershman, S.J. (2017). Compositional inductive biases in function learning. *Cognitive Psychology, 99,* 44-79.
159. Gershman, S.J. (2017). Dopamine, inference and uncertainty. *Neural Computation.*

160. Gershman, S.J., Zhou, J., & Kommers, C. (2017). Imaginative reinforcement learning: computational principles and neural mechanisms. *Journal of Cognitive Neuroscience*, *29*, 2103-2113.
161. Gershman, S.J. & Beck, J.M. (2017). Complex probabilistic inference: from cognition to neural computation. In A. Moustafa (Ed.) *Computational Models of Brain and Behavior*. Wiley-Blackwell.
162. Momennejad, I., Russek, E., Cheong, J.H., Botvinick, M.M., Daw, N.D., & Gershman, S.J. (2017). The successor representation in human reinforcement learning. *Nature Human Behaviour, 1,* 680-692.
163. Linderman, S.W., & Gershman, S.J. (2017). Using computational theory to constrain statistical models of neural data. *Current Opinion in Neurobiology*, 46, *14-24*.
164. Saeedi, A., Kulkarni, T., Mansinghka, V.K., & Gershman, S.J. (2017). Variational particle approximations. *Journal of Machine Learning Research, 18*, 1-29.
165. Kool, W., Gershman, S.J., & Cushman, F.A. (2017). Cost-benefit arbitration between multiple reinforcement learning systems. *Psychological Science, 28,* 1321-1333.
166. Dasgupta, I., Schulz, E., & Gershman, S.J. (2017). Where do hypotheses come from? *Cognitive Psychology, 96,* 1-25.
167. Dasgupta, I., Schulz, E., Goodman, N.D., & Gershman, S.J. (2017). Amortized hypothesis generation. *Proceedings of the 39th Annual Conference of the Cognitive Science Society*.
168. Russek, E., Momennejad, I., Botvinick, M.M., Gershman, S.J., & Daw, N.D. (2017). Predictive representations can link model-based reinforcement learning to model-free mechanisms. *PLOS Computational Biology, 13,* e1005768.
169. Starkweather, C.K., Babayan, B.M., Uchida, N., & Gershman, S.J. (2017). Dopamine reward prediction errors reflect hidden state inference across time. *Nature Neuroscience, 20,* 581-589.
170. Gershman, S.J. (2017). Predicting the past, remembering the future. *Current Opinion in Behavioral Sciences, 17,* 7-13.
171. Thaker, P., Tenenbaum, J.B., & Gershman, S.J. (2017). Online learning of symbolic concepts. *Journal of Mathematical Psychology, 77, 10-20.*
172. Gershman, S.J., Pouncy, H.T., & Gweon, H. (2017). Learning the structure of social influence. *Cognitive Science, 41, 545-575.*
173. Gershman, S.J. (2017). Context-dependent learning and causal structure. *Psychonomic Bulletin & Review, 24*, 557-565.
174. Gershman, S.J., Malmaud, J., & Tenenbaum, J.B. (2017). Structured representations of utility in combinatorial domains. *Decision, 4, 67-86.*
175. Gershman, S.J. (2017). Reinforcement learning and causal models. In M. Waldmann, Ed, *Oxford Handbook of Causal Reasoning*. Oxford University Press.
176. Gershman, S.J. & Daw, N.D. (2017). Reinforcement learning and episodic memory in humans and animals: an integrative framework. *Annual Review of Psychology, 68,* 101-128.
177. Gershman, S.J., Monfils, M.-H., Norman, K.A., & Niv, Y. (2017). The computational nature of memory modification. *eLife*, *6*, e23763.

178. Gershman, S.J. (2017). On the blessing of abstraction. *The Quarterly Journal of Experimental Psychology*, *70,* 361-365.
179. Gershman, S.J., Tenenbaum, J.B., & Jäkel, F.J. (2016). Discovering hierarchical motion structure. *Vision Research, 126,* 232-241.
180. Pereira, F., Gershman, S.J., Ritter, S., & Botvinick, M.M. (2016). A comparative evaluation of off-the-shelf distributed semantic representations for modelling behavioural data. *Cognitive Neuropsychology, 33*, 175-190.
181. Schulz, E., Tenenbaum, J.B., Duvenaud, D., Speekenbrink, M., & Gershman, S.J. (2016). Probing the compositionality of intuitive functions. *Advances in Neural Information Processing Systems, 29*.
182. Gershman, S.J., Gerstenberg, T., Baker, C.L., & Cushman, F.A. (2016). Plans, habits, and theory of mind. *PLOS One, 11*, e0162246.
183. Kool, W., Cushman, F.A., & Gershman, S.J. (2016). When does model-based control pay off? *PLOS Computational Biology, 12,* e1005090.
184. Ullman, T.D., Siegel, M., Tenenbaum, J.B., & Gershman, S.J. (2016). Coalescing the vapors of human experience into a viable and meaningful comprehension. *Proceedings of the 38th Annual Conference of the Cognitive Science Society*.
185. Batmanghelich, K., Saeedi, A., Narasimhan, K., & Gershman, S.J. (2016). Nonparametric spherical topic modeling with word embeddings. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*.
186. Gershman, S.J. (2016). Empirical priors for reinforcement learning models. *Journal of Mathematical Psychology, 71,* 1-6.
187. Tervo, D.G.R., Tenenbaum, J.B., & Gershman, S.J. (2016). Towards the neural implementation of structure learning. *Current Opinion in Neurobiology, 37*, 99-105.
188. Gershman, S.J. (2015). A unifying probabilistic view of associative learning. *PLOS Computational Biology, 11,* e1004567.
189. Gershman, S.J. (2015). Do learning rates adapt to the distribution of rewards? *Psychonomic Bulletin & Review, 22,* 1320-1327.
190. Gershman, S.J., Norman, K.A., & Niv, Y. (2015). Discovering latent causes in reinforcement learning. *Current Opinion in Behavioral Sciences, 5,* 43-50.
191. Gershman, S.J. & Tenenbaum, J.B. (2015). Phrase similarity in humans and machines. *Proceedings of the 37th Annual Conference of the Cognitive Science Society*.
192. Schulz, E., Tenenbaum, J.B., Reshef, D.N., Speekenbrink, M., & Gershman, S.J. (2015). Assessing the perceived predictability of functions. *Proceedings of the 37th Annual Conference of the Cognitive Science Society*.
193. Gershman, S.J., Horvitz, E.J., & Tenenbaum, J.B. (2015). Computational rationality: a converging paradigm for intelligence in brains, minds and machines. *Science, 349,* 273-278.
194. Gershman, S.J. & Hartley, C.A. (2015). Individual differences in learning predict the return of fear. *Learning & Behavior, 43,* 243-250.

195. Niv, Y., Daniel, R., Geana, A., Gershman, S.J., Leong, Y.C., Radulescu, A., & Wilson, R.C. (2015). Reinforcement learning in multidimensional environments relies on attention mechanisms. *Journal of Neuroscience, 35, 8145-8157.*
196. Huys, Q.J.M., Lally, N., Faulkner, P., Eshel, N., Seifritz, E., Gershman, S.J., Dayan, P., & Roiser, J.P. (2015). The interplay of approximate planning strategies. *Proceedings of the National Academy of Sciences, 112*, 3098-3103.
197. Gershman, S.J. & Niv, Y. (2015). Novelty and inductive generalization in human reinforcement learning. *Topics in Cognitive Science,* 1-25.
198. Gershman, S.J., Frazier, P.I., & Blei, D.M. (2015). Distance dependent infinite latent feature models. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 37,* 334-345.
199. Austerweil, J.L., Gershman, S.J., Tenenbaum, J.B., & Griffiths, T.L. (2015). Structure and flexibility in Bayesian models of cognition. In J.R. Busemeyer, J.T. Townsend, Z. Wang, & A. Eidels, Eds, *Oxford Handbook of Computational and Mathematical Psychology*. Oxford University Press.
200. Stachenfeld, K.L., Botvinick, M.M., & Gershman, S.J. (2014). Design principles of the hippocampal cognitive map. *Advances in Neural Information Processing Systems 27*.
201. Gershman, S.J., Radulescu, A., Norman, K.A., & Niv, Y. (2104). Statistical computations underlying the dynamics of memory updating. *PLOS Computational Biology, 10,* e1003939.
202. Gershman, S.J., Blei, D.M., Norman, K.A., & Sederberg, P.B. (2014). Decomposing spatiotemporal brain images into topographic latent sources. *NeuroImage, 98,* 91-102.
203. Gershman, S.J. (2014). The penumbra of learning: A statistical theory of synaptic tagging and capture. *Network: Computation in Neural Systems, 25,* 97-115.
204. Soto, F.A., Gershman, S.J., & Niv, Y. (2014). Explaining compound generalization in associative and causal learning through rational principles of dimensional generalization. *Psychological Review, 121,* 526-558.
205. Gershman, S.J. & Goodman, N.D. (2014). Amortized inference in probabilistic reasoning. *Proceedings of the 36th Annual Conference of the Cognitive Science Society.*
206. Tsividis, P., Gershman, S.J., Tenenbaum, J.B., & Schulz, L. (2014). Information selection in noisy environments with large action spaces. *Proceedings of the 36th Annual Conference of the Cognitive Science Society.*
207. Feng, S.F., Schwemmer, M., Gershman, S.J., & Cohen, J.D. (2014). Multitasking vs. multiplexing: Toward a normative account of capacity constraints in cognitive control. *Cognitive, Affective, and Behavioral Neuroscience*, *14,* 129-146.
208. Gershman, S.J., Moustafa, A.A., & Ludvig, E.A. (2014). Time representation in reinforcement learning models of the basal ganglia. *Frontiers in Computational Neuroscience.*
209. Gershman, S.J., Markman, A.B., & Otto, A.R. (2014). Retrospective revaluation in sequential decision making: a tale of two systems. *Journal of Experimental Psychology: General, 143,* 182-194.
210. Gershman, S.J. (2014). Dopamine ramps are a consequence of reward prediction errors. *Neural Computation, 26,* 467-471.

211. Gershman, S.J. (2013). Computation with dopaminergic modulation. In Jaeger, D., Jung, R. (Ed.) *Encyclopedia of Computational Neuroscience*. Springer.
212. Gershman, S.J. (2013). Bayesian behavioral data analysis. In Jaeger, D., Jung, R. (Ed.) *Encyclopedia of Computational Neuroscience*. Springer.
213. Gershman, S.J., Jones, C.E., Norman, K.A., Monfils, M.-H., & Niv, Y. (2013). Gradual extinction prevents the return of fear. *Frontiers in Behavioral Neuroscience.* doi: 10.3389/fnbeh.2013.00164.
214. Detre, G.J., Natarajan, A., Gershman, S.J., & Norman, K.A. (2013). Moderate levels of activation lead to forgetting in the think/no-think paradigm. *Neuropsychologia*, *51,* 2371-2388.
215. Gershman, S.J., & Niv, Y. (2013). Perceptual estimation obeys Occam's razor. *Frontiers in Psychology, 23,* doi:10.3389/fpsyg.2013.00623.
216. Christakou, A., Gershman, S.J., Niv, Y., Simmons, A., Brammer, M., & Rubia, K. (2013). Neural and psychological maturation of decision-making in adolescence and young adulthood. *Journal of Cognitive Neuroscience, 25*, 1807-1823.
217. Gershman, S.J, Schapiro, A.C., Hupbach, A., & Norman, K.A. (2013). Neural context reinstatement predicts memory misattribution. *Journal of Neuroscience, 33,* 8590-8595.
218. Otto, A.R., Gershman, S.J., Markman, A.B., & Daw, N.D. (2013). The curse of planning: Dissecting multiple reinforcement learning systems by taxing the central executive. *Psychological Science*, *24,* 751-761.
219. Wingate, D., Diuk, C., O'Donnell, T.J., Tenenbaum, J.B., Gershman, S.J. (2013). Compositional policy priors. *CSAIL Technical Report 2013-007.*
220. Gershman, S.J., Jäkel, F.J., & Tenenbaum, J.B. (2013). Bayesian vector analysis and the perception of hierarchical motion. *Proceedings of the 35th Annual Conference of the Cognitive Science Society*.
221. Gershman, S.J. & Niv, Y (2012). Exploring a latent cause model of classical conditioning. *Learning & Behavior, 40*, 255-268.
222. Gershman, S.J., Hoffman, M.D., & Blei, D.M. (2012). Nonparametric variational inference. *Proceedings of the 29th International Conference on Machine Learning*.
223. Gershman, S.J., Moore, C.D., Todd, M.T., Norman, K.A., & Sederberg, P.B. (2012). The successor representation and temporal context. *Neural Computation, 5*, 1553-1568.
224. Gershman, S.J. & Blei, D.M. (2012). A tutorial on Bayesian nonparametric models. *Journal of Mathematical Psychology, 56*, 1-12.
225. Gershman, S.J. & Daw, N.D. (2012). Perception, action and utility: the tangled skein. In M. Rabinovich, K. Friston, P. Varona (Ed.) *Principles of Brain Dynamics: Global State Interactions*. MIT Press.
226. Gershman, S.J., Vul, E., & Tenenbaum, J.B. (2012). Multistability and perceptual inference. *Neural Computation, 24*, 1-24.
227. Gershman, S.J., Blei, D.M., Pereira, F., & Norman, K.A. (2011). A topographic latent source model for fMRI data. *NeuroImage, 57*, 89-100.

228. Sederberg, P.B., Gershman, S.J., Polyn, S.M., & Norman, K.A. (2011). Human memory reconsolidation can be explained using the Temporal Context Model. *Psychonomic Bulletin and Review, 18*, 455-468.
229. Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P., & Dolan, R.J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron, 69*, 1204-1215.
230. Gershman, S.J. & Wilson, R.C. (2010). The neural costs of optimal control, *Advances in Neural Information Processing Systems 23*.
231. Gershman, S.J, Cohen, J.D., & Niv, Y. (2010). Learning to selectively attend, *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*.
232. Gershman, S.J & Niv, Y. (2010). Learning latent structure: Carving nature at its joints, *Current Opinion in Neurobiology, 20*, 1-6.
233. Gershman, S.J., Blei, D.M., & Niv, Y. (2010). Context, learning, and extinction, *Psychological Review, 117*, 197-209.
234. Gershman, S.J., Pesaran, B., & Daw, N.D. (2009). Human reinforcement learning subdivides structured action spaces by learning effector-specific values, *Journal of Neuroscience, 29*, 13524-13531.
235. Gershman, S.J., Vul, E., & Tenenbaum, J.B. (2009). Perceptual multistability as Markov chain Monte Carlo inference, *Advances in Neural Information Processing Systems 22*.
236. Socher, R., Gershman, S.J., Perotte, A., Sederberg, P.B., Blei, D.M., & Norman, K.A. (2009). A Bayesian analysis of dynamics in free recall, *Advances in Neural Information Processing Systems 22*.

EDITORIAL POSITIONS

Editor-in-chief: Open Mind
Special issue guest editor: Current Opinion in Behavioral Sciences (special issue on artificial intelligence, 2019), Topics in Cognitive Science (special issue on computational approaches to social cognition, 2019)

Former editorial positions

Deputy editor: PLOS Computational Biology

Associate editor: Cognitive Science

Ad hoc editor: eLife

Reviewing editor: Psychological Review