

Curriculum Vitae

Samuel J. Gershman

PERSONAL DETAILS

Samuel J. Gershman
Department of Psychology and Center for Brain Science
Harvard University
Northwest Building, room 295.05
52 Oxford St.
Cambridge, MA 02138
gershman@fas.harvard.edu
773-607-9817

EMPLOYMENT HISTORY

Assistant Professor	2015-
Department of Psychology and Center for Brain Science, Harvard University	
Postdoctoral fellow	2013-2015
Department of Brain and Cognitive Sciences, MIT	

EDUCATION

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

FUNDING

2017-2020	ONR Science of Autonomy – Structured reinforcement learning in the brain (role: PI)
2017-2022	ONR MURI - A Computational Cognitive Neuroscience Approach to Understanding Event Representation and Episodic Memory (role: co-I)
2015-2018	NIH R01 - Representational foundations of adaptive behavior in natural and artificial agents (role: co-I with Nathaniel Daw (PI) and George Konidaris)
2017-2018	Harvard Foundations of Human Behavior Initiative - Computational and Neural Mechanisms of Information Seeking Behavior (role: PI)
2016-2017	Harvard Foundations of Human Behavior Initiative - Goals, habits and theory of mind: experimental and computational studies (role: co-PI with Fiery Cushman)
2016-2017	Harvard Brain Initiative Collaborative Seed Grant - Neurocomputational mechanisms of structure discovery (role: co-PI with Jan Drugowitsch)

- 2016- Google Research Grant - Research focusing on the use of video games environments to probe human planning, exploration, and inference (role: PI)
- 2015-2016 Harvard Mind/Brain/Behavior Initiative - Representation of hidden state in the dopamine system (role: co-PI with Naoshige Uchida)

FELLOWSHIPS and AWARDS

- 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

INVITED TALKS

- Cognitive lunch seminar, Indiana University (April, 2010)
- Pavlovian Society annual meeting, Baltimore (October, 2010)
- Cognitive science seminar, University of Texas (February 2011)
- Organization for Human Brain Mapping meeting (June 2012)
- Oxford University, Department of Experimental Psychology (July 2012)
- Gatsby Computational Neuroscience Unit (July 2012)
- Pavlovian Society annual meeting, Jersey City (September, 2012)
- Society for computational modeling of associative learning, Jersey City (September, 2012)
- Brown University, Department of Cognitive, Linguistic and Psychological Sciences colloquium (December, 2012)
- Harvard University, Center for Brain Sciences (February 2014)
- University of California, Berkeley, Department of Psychology (February 2014)
- Columbia University, Department of Psychology (February 2014)
- University of Zurich, Neuroeconomics seminar (September 2014)
- Tufts University, Cognitive science seminar (October 2014)
- Boston University, Brain, Behavior and Cognition seminar (March 2015)
- Caltech, Neuroeconomics seminar (April 2015)
- Harvard Center for Brain/Mind seminar series (June 2015)
- Yale University, Current Works in Behavior, Genetics, and Neuroscience (February 2016)
- University of Massachusetts, Amherst, Cognitive science seminar (September 2016)

- Google DeepMind (October 2016)
- Boston Veteran's Affairs hospital, Neuroimaging/neuropsychology lecture series (October 2016)
- University of Pennsylvania, Computational Neuroscience seminar (November 2016)
- Champalimaud Center for the Unknown, Quantitative Neuroscience Seminar (January 2017)
- UC Davis, Neuroscience seminar (March 2017)
- New York University, Neuroeconomics colloquium (April 2017)
- Reinforcement learning and decision making conference, keynote speaker (June 2017)
- Sloan-Norris Workshop on Attentional and Perceptual Foundations of Behavior, New York, NY (October, 2017)
- Brown University, Beyond Deep Learning workshop (January 2018)
- Stanford University, Mind, Brain and Computation seminar (January 2018)
- Gatsby Computational Neuroscience Unit, UCL, External Seminar (March 2018)
- Functional Imaging Laboratory, UCL, special seminar (March 2018)
- University of Durham, Probabilistic Brain Workshop, keynote speaker (March 2018)
- University of Oxford, Neurotheory seminar (March 2018)
- Tufts University, Predictive processing symposium (April 2018)

MANUSCRIPTS UNDER REVIEW

1. Chang, L.W., Gershman, S.J., & Cikara, M. (submitted). Divisive normalization drives context dependence in social choice.
2. Tomov, M.S., Dorfman, H.M., & Gershman, S.J. (submitted). Neural computations underlying causal structure learning.
3. Babayan, B.M., Uchida, N., & Gershman, S.J. (submitted). Belief state representation in the dopamine system.
4. Leon-Villagra, P., Schulz, E., Speekenbrink, M., Gershman, S.J., & Lucas, C.G. (submitted). One-shot compositional function learning.
5. Christakou, A., Gershman, S.J., Simmons, A., Murphy, C., Giampetro, V., Brammer, M., & Rubia, K. (submitted). Atypical maturation of decision-making in adolescent males with ADHD.
6. Lau, T., Pouncy, H.T., Gershman, S.J., & Cikara, M. (submitted). Discovering social groups via latent structure learning.
7. Gershman, S.J. (submitted). How to never be wrong.

PUBLICATIONS

1. Starkweather, C.K., Gershman, S.J., & Uchida, N. (in press). Medial prefrontal cortex shapes dopamine reward prediction errors under state uncertainty. *Neuron*.
2. Millner, A.J., Gershman, S.J., Nock, M.K., & Ouden, H.D. (in press). Pavlovian control of escape and avoidance. *Journal of Cognitive Neuroscience*.

3. Gershman, S.J. (2018). Deconstructing the human algorithms for exploration. *Cognition*, *173*, 34-42.
4. 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.
5. 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.
6. 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.
7. Stachenfeld, K.L., Botvinick, M.M., & Gershman, S.J. (2017). The hippocampus as a predictive map. *Nature Neuroscience*, *20*, 1643-1653.
8. 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.
9. Schulz, E., Tenenbaum, J.B., Duvenaud, D., Speekenbrink, M., & Gershman, S.J. (2017). Compositional inductive biases in function learning. *Cognitive Psychology*, *99*, 44-79.
10. Gershman, S.J. (2017). Dopamine, inference and uncertainty. *Neural Computation*.
11. Gershman, S.J., Zhou, J., & Kommer, C. (2017). Imaginative reinforcement learning: computational principles and neural mechanisms. *Journal of Cognitive Neuroscience*, *29*, 2103-2113.
12. 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.
13. 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.
14. Linderman, S.W., & Gershman, S.J. (2017). Using computational theory to constrain statistical models of neural data. *Current Opinion in Neurobiology*, *46*, 14-24.
15. Saeedi, A., Kulkarni, T., Mansinghka, V.K., & Gershman, S.J. (2017). Variational particle approximations. *Journal of Machine Learning Research*, *18*, 1-29.
16. Kool, W., Gershman, S.J., & Cushman, F.A. (2017). Cost-benefit arbitration between multiple reinforcement learning systems. *Psychological Science*, *28*, 1321-1333.
17. Dasgupta, I., Schulz, E., & Gershman, S.J. (2017). Where do hypotheses come from? *Cognitive Psychology*, *96*, 1-25.
18. 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*.

19. 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.
20. 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.
21. Gershman, S.J. (2017). Predicting the past, remembering the future. *Current Opinion in Behavioral Sciences*, *17*, 7-13.
22. Thaker, P., Tenenbaum, J.B., & Gershman, S.J. (2017). Online learning of symbolic concepts. *Journal of Mathematical Psychology*, *77*, 10-20.
23. Gershman, S.J., Pouncy, H.T., & Gweon, H. (2017). Learning the structure of social influence. *Cognitive Science*, *41*, 545-575.
24. Gershman, S.J. (2017). Context-dependent learning and causal structure. *Psychonomic Bulletin & Review*, *24*, 557-565.
25. Gershman, S.J., Malmaud, J., & Tenenbaum, J.B. (2017). Structured representations of utility in combinatorial domains. *Decision*, *4*, 67-86.
26. Gershman, S.J. (2017). Reinforcement learning and causal models. In M. Waldmann, Ed, *Oxford Handbook of Causal Reasoning*. Oxford University Press.
27. 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.
28. Gershman, S.J., Monfils, M.-H., Norman, K.A., & Niv, Y. (2017). The computational nature of memory modification. *eLife*.
29. Gershman, S.J. (2017). On the blessing of abstraction. *The Quarterly Journal of Experimental Psychology*, *70*, 361-365.
30. Gershman, S.J., Tenenbaum, J.B., & Jäkel, F.J. (2016). Discovering hierarchical motion structure. *Vision Research*, *126*, 232-241.
31. 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.
32. 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*.
33. Gershman, S.J., Gerstenberg, T., Baker, C.L., & Cushman, F.A. (2016). Plans, habits, and theory of mind. *PLOS One*, *11*, e0162246.
34. Kool, W., Cushman, F.A., & Gershman, S.J. (2016). When does model-based control pay off? *PLOS Computational Biology*, *12*, e1005090.
35. 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*.
36. 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*.

37. Gershman, S.J. (2016). Empirical priors for reinforcement learning models. *Journal of Mathematical Psychology*, *71*, 1-6.
38. 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.
39. Gershman, S.J. (2015). A unifying probabilistic view of associative learning. *PLOS Computational Biology*, *11*, e1004567.
40. Gershman, S.J. (2015). Do learning rates adapt to the distribution of rewards? *Psychonomic Bulletin & Review*, *22*, 1320-1327.
41. Gershman, S.J., Norman, K.A., & Niv, Y. (2015). Discovering latent causes in reinforcement learning. *Current Opinion in Behavioral Sciences*, *5*, 43-50.
42. Gershman, S.J. & Tenenbaum, J.B. (2015). Phrase similarity in humans and machines. *Proceedings of the 37th Annual Conference of the Cognitive Science Society*.
43. 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*.
44. 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.
45. Gershman, S.J. & Hartley, C.A. (2015). Individual differences in learning predict the return of fear. *Learning & Behavior*, *43*, 243-250.
46. 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.
47. 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.
48. Gershman, S.J. & Niv, Y. (2015). Novelty and inductive generalization in human reinforcement learning. *Topics in Cognitive Science*, 1-25.
49. 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.
50. 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.
51. Stachenfeld, K.L., Botvinick, M.M., & Gershman, S.J. (2014). Design principles of the hippocampal cognitive map. *Advances in Neural Information Processing Systems* *27*.
52. Gershman, S.J., Radulescu, A., Norman, K.A., & Niv, Y. (2014). Statistical computations underlying the dynamics of memory updating. *PLOS Computational Biology*, *10*, e1003939.
53. 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.

54. Gershman, S.J. (2014). The penumbra of learning: A statistical theory of synaptic tagging and capture. *Network: Computation in Neural Systems*, 25, 97-115.
55. 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.
56. Gershman, S.J. & Goodman, N.D. (2014). Amortized inference in probabilistic reasoning. *Proceedings of the 36th Annual Conference of the Cognitive Science Society*.
57. 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*.
58. 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.
59. Gershman, S.J., Moustafa, A.A., & Ludvig, E.A. (2014). Time representation in reinforcement learning models of the basal ganglia. *Frontiers in Computational Neuroscience*.
60. 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.
61. Gershman, S.J. (2014). Dopamine ramps are a consequence of reward prediction errors. *Neural Computation*, 26, 467-471.
62. Gershman, S.J. (2013). Computation with dopaminergic modulation. In Jaeger, D., Jung, R. (Ed.) *Encyclopedia of Computational Neuroscience*. Springer.
63. Gershman, S.J. (2013). Bayesian behavioral data analysis. In Jaeger, D., Jung, R. (Ed.) *Encyclopedia of Computational Neuroscience*. Springer.
64. 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.
65. 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.
66. Gershman, S.J., & Niv, Y. (2013). Perceptual estimation obeys Occam's razor. *Frontiers in Psychology*, 23, doi:10.3389/fpsyg.2013.00623.
67. 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.
68. Gershman, S.J., Schapiro, A.C., Hupbach, A., & Norman, K.A. (2013). Neural context reinstatement predicts memory misattribution. *Journal of Neuroscience*, 33, 8590-8595.
69. 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.
70. Wingate, D., Diuk, C., O'Donnell, T.J., Tenenbaum, J.B., Gershman, S.J. (2013). Compositional policy priors. *CSAIL Technical Report 2013-007*.

71. 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*.
72. Gershman, S.J. & Niv, Y (2012). Exploring a latent cause model of classical conditioning. *Learning & Behavior, 40*, 255-268.
73. Gershman, S.J., Hoffman, M.D., & Blei, D.M. (2012). Nonparametric variational inference. *Proceedings of the 29th International Conference on Machine Learning*.
74. 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.
75. Gershman, S.J. & Blei, D.M. (2012). A tutorial on Bayesian nonparametric models. *Journal of Mathematical Psychology, 56*, 1-12.
76. 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.
77. Gershman, S.J., Vul, E., & Tenenbaum, J.B. (2012). Multistability and perceptual inference. *Neural Computation, 24*, 1-24.
78. Gershman, S.J., Blei, D.M., Pereira, F., & Norman, K.A. (2011). A topographic latent source model for fMRI data. *NeuroImage, 57*, 89-100.
79. 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.
80. 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.
81. Gershman, S.J. & Wilson, R.C. (2010). The neural costs of optimal control, *Advances in Neural Information Processing Systems 23*.
82. Gershman, S.J., Cohen, J.D., & Niv, Y. (2010). Learning to selectively attend, *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*.
83. Gershman, S.J. & Niv, Y. (2010). Learning latent structure: Carving nature at its joints, *Current Opinion in Neurobiology, 20*, 1-6.
84. Gershman, S.J., Blei, D.M., & Niv, Y. (2010). Context, learning, and extinction, *Psychological Review, 117*, 197-209.
85. 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.
86. Gershman, S.J., Vul, E., & Tenenbaum, J.B. (2009). Perceptual multistability as Markov chain Monte Carlo inference, *Advances in Neural Information Processing Systems 22*.
87. 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*.

AD HOC REVIEWER (alphabetical order)

Biological Cybernetics, Biological Psychiatry, Brain & Cognition, Cognition, Cognitive Science, CABN, Current Biology, Decision, Frontiers in Decision Neuroscience, Journal of Cognitive Neuroscience, Journal of Experimental Psychology: Learning, Memory & Cognition, Journal of Experimental Psychology: General, Journal of Machine Learning Research, Journal of Mathematical Psychology, Journal of Neuroscience, Journal of Neuroscience Methods, Learning & Memory, Neural Computation, Nature Communications, Nature Neuroscience, Neurobiology of Aging, NeuroImage, Neuron, Neuropsychologia, PLOS Computational Biology, PNAS, Psychological Review, Psychonomic Bulletin & Review, Science

ASSOCIATE EDITOR

PLOS Computational Biology

AD HOC EDITOR

eLife

REVIEWING EDITOR

Psychological Review

TEACHING

Cognition: Mind and Brain (Spring 2007), Columbia University—teaching assistant

Animal learning and decision making: psychological, computational and neural perspectives (Fall 2010), Princeton University—teaching assistant

Computational Cognitive Neuroscience (Fall 2016, yearly thereafter), Harvard University

Debugging the Brain (Spring 2018, yearly thereafter), Harvard University