Curriculum Vitae
Seth A. Herd

Education:

Ph.D., Psychology & Neuroscience, University of Colorado at Boulder, 2005
Thesis: "Mechanisms of visual search."

M.A., Cognitive Psychology, University of Colorado at Boulder, 2000

B.A., Physics, Earlham College, Richmond IN, 1997
"Honors" (highest possible commendation)

Professional Experience:

Principal Investigator, AI safety project, Theiss Research – Sept. 2015 – Sept. 2018

Wrote proposal and performed research for *Stability of Neuromorphic Motivational Systems*, Future of Life Institute. Performed background research on safety issues in artificial general intelligence; adapted current theories of human motivation and brain function to AGI safety.

Project Manager and Senior Modeler – eCortex, Inc. October 2012 – present

Managed and performed research on government-funded research contracts: *The Power and Pitfalls of Flexible Human Cognitive Control*, ONR, $2m total; *Neural Mechanisms of Adaptive Human Executive Control*, ONR, $2m total; *Neural Mechanisms of Risky Decision-Making*, subcontract to USC, NIH funding, $650k total; *Automated Archerfish EO / Sonar Detection Techniques for Naval Mines,* NAVSEA, $1.5m total. Planned project approaches, goals, and milestones; hired personnel, managed task divisions and goals, planned and monitored technical progress on daily/weekly basis; designed and implemented models; performed secondary research and synthesized theories.

CEO & General Manager, eCortex, Inc., September 2011 – present

Managed all aspects of company business, including managing administration of ongoing projects, making strategy for and writing new proposals, and working with board of directors on long-term business plans. Supervised support staff, programming staff, and administrative staff. As CEO since Jan. 2013, additionally assumed responsibility for and supervision of all aspects of the business.

Senior Research Associate, University of Colorado, January 2013 – October 2015

Assisted lab head in planning and managing student and postdoctoral projects on projects including fMRI and modeling of neural mechanisms of executive function, motivation, vision, and general cortical processing.

Principal Investigator, ICArUS project, eCortex, Inc., August 2011 - December 2014

Managed company participation as a subcontractor to HRL industries, Inc., in multi-team IARPA-funded effort to create integrated, neutrally realistic models of how many brain areas work together to perform geospatial intelligence sensemaking tasks. Coordinated company efforts, and collaborated with HRL and University of Colorado PI's in designing integrated model, perhaps the most ambitious neural network model of brain function to date.

Consultant, eCortex, IARPA seedling on embodied cognition, June 2010 - May 2011

Served as a neural network modeling and theory of visual cognition consultant. Designed models to differentiate figure from ground in complex environments using learning signals derived from binocular vision.

Staff Modeler, eCortex, Inc., Navy SBIR on artificial vision, Jan. 2007- Dec. 2007

Developed and tested LVIS neural network model of human visual system on task of detecting floating mines in video footage of ocean waves.

Professional Research Assistant, University of Colorado, 2006 - 2012

Planned, performed, and published research on neural network theories of human cognitive function on topics including human executive function, visual attention, object recognition, cognitive sequencing, strategic decision-making, task switching, working memory, and language learning. Supervised interns and directed graduate student projects. Assisted in coordinating all aspects of O'Reilly lab function, including planning and coordinating grant proposals and long-term projects.

Instructor, University of Colorado, 2005, 2006, 2007, 2009

Taught classes "introduction to cognitive psychology" and "Computational Cognitive Neuroscience". Achieved high marks in student evaluations by clearly explaining computational cognitive neuroscience and its potential importance.

Peer Reviewed Publications:

Herd, S., Krueger, K., Nair, A., Mollick, J., & O'Reilly, R. (2021). Neural mechanisms of human decision-making. *Cognitive, Affective, & Behavioral Neuroscience*, *21*(1), 35-57.

Mollick, J. A., Hazy, T. E., Krueger, K. A., Nair, A., Mackie, P., Herd, S. A., & O'Reilly, R. C. (2020). A systems-neuroscience model of phasic dopamine. *Psychological Review*, *127*(6), 972.

O'Reilly, R. C., Nair, A., Russin, J. L., & Herd, S. A. (2020). How sequential interactive processing within frontostriatal loops supports a continuum of habitual to controlled processing. *Frontiers in psychology*, *11*, 380.

O'Reilly, R. C., Russin, J., & Herd, S. A. (2019). Computational models of motivated frontal function. *Handbook of clinical neurology*, *163*, 317-332.

Jilk, D. J., Herd, S., Read, S. J., & O'Reilly, R. C. (2017). Anthropomorphic reasoning about neuromorphic AGI safety. *Journal of Experimental & Theoretical Artificial Intelligence*, *29*(6), 1337-1351.

Herd, S. A., Hazy, T. E., Chatham, C. H., Brant, A. M., & Friedman, N. P. (2014). A neural network model of individual differences in task switching abilities. *Neuropsychologia*, 62, 375-389.

Herd, S., Szabados, A., Vinokurov, Y., Lebiere, C., Cline, A & O'Reilly, R.C. (2014) Integrating theories of motor sequencing in the SAL hybrid architecture. *Biologically Inspired Cognitive Architectures* 8, 100-108

Vinokurov, Y., Lebiere, C., Szabados, A., Herd, S.A., & O'Reilly, R.C. (2013). Integrating top-down expectations with bottom-up perceptual processing in a hybrid neural-symbolic architecture. *Biologically Inspired Cognitive Architectures*, *6*, 140-146.

Herd, S.A., Kruger, K., Huang, T.R., Kriete, T., Hazy, T.E., & O'Reilly, R.O. (2013). Strategic cognitive sequencing: a computational cognitive neuroscience approach. *Computational Intelligence and Neuroscience*, Volume 2013, Article ID 149329

Huang, T.R., Herd, S.A., Hazy, T.E., & O'Reilly, R.O. (2013). Assembling old tricks for new tasks: a neural model of instructional learning and control. *Journal of Cognitive Neuroscience, 25*(6), 843-851

O'Reilly, R.C., Wyatte, D., Herd, S., Mingus, B., & Jilk, D. (2013). Recurrent processing during object recognition. *Frontiers in Psychology, 4*(124)

Wyatte, D., Herd, S.A., Mingus, B., O'Reilly, R.C. (2012). The Role of Competitive Inhibition and Top-Down Feedback in Binding during Object Recognition. *Frontiers in Psychology, 3*(182)

Mingus, B., Kriete, T., Herd, S., Wyatte, D., Latimer, K., O'Reilly, R.C. (2011) Generalization of Figure-Ground Segmentation from Binocular to Monocular Vision in an Embodied Biological Brain Model. *Proceedings of the 4th Conference on Artificial General Intelligence: Lecture Notes in Computer Science 6830*, 351-356

O'Reilly, R.C., Pauli, W. & Herd, S.A. (2011). Computational models of cognitive control. *Current Opinion in Neurobiology*. 20(2), 257-261

Munakata Y, Herd S.A., Chatham C.H., Depue B.E., Banich M.T., O'Reilly R.C. (2011) A unified framework for inhibitory control. *Trends in Cognitive Science;15*(10), 453-459

Herd, S.A., Urland, G., Mingus, B., O'Reilly, R.C. (2011) Human-artificial-intelligence hybrid learning systems. *Frontiers in Artificial Intelligence and Applications 233*, 132 - 137

Vinokurov, Y., Lebiere, C., Herd, S., & O'Reilly, R. (2011). A metacognitive classifier using a hybrid act-r/leabra architecture. *Lifelong Learning: Papers from the 2011 AAAI Workshop (WS-11-15)*.

Chatham, C.H., Herd, S.A., Brant, A.M., Hazy, T.E., Miyake, A., O'Reilly, R.C. & Friedman, N.P. (2011) From the executive network to the executive functions: A computational model of the *n*-back task. Journal of Cognitive Neuroscience, 23(11), 3598-3619

Herd, S.A., Mingus, B. & O'Reilly, R.C. (2010). Dopamine and self-directed learning. Proceeding of the 2010 conference on Biologically Inspired Cognitive Architectures 2010. Amsterdam: IOS Press.

O'Reilly, R. C., Herd, S. A., & Pauli, W. M. (2010). Computational models of cognitive control. *Current opinion in neurobiology, 20*(2), 257-261.

Taatgen, N.A., Juvina, I., Herd, S.A., Jilk, D. & Martens, S. (2007). Attentional Blink: an internal traffic jam? *Proceedings of the eighth International Conference on Cognitive Modeling* (pp. 91-96). New York: Psychology Press.

Herd, S.A., Banich, M.T. & O'Reilly, R.C. (2006). Neural mechanisms of cognitive control: an integrative model of Stroop task performance. *Journal of Cognitive Neuroscience, 18*(1), 22-32

Herd, S.A., & O'Reilly, R.C. (2005). Parallel models of visual search. *Vision Research, 45*(24). 2987-2992

Book Chapters & Preprints

Herd, S.A., Read, S.J., O'Reilly, R.C., & Jilk, D.J. (2018). Goal Changes in Intelligent Agents. In Yampolskiy, Y., (Ed) *Artificial intelligence safety and security.*

O'Reilly, R. C., Hazy, T. E., & Herd, S. A. (2016). The leabra cognitive architecture: how to play 20 principles with nature and win!. *The Oxford Handbook of Cognitive Science*.

O'Reilly, R. C., Hazy, T. E., Mollick, J., Mackie, P., & Herd, S. (2014). Goal-driven cognition in the brain: a computational framework. *arXiv preprint arXiv:1404.7591*.

O'Reilly, R. C., Petrov, A. A., Cohen, J. D., Lebiere, C. J., Herd, S. A., Kriete, T., ... & Symons, J. (2013). How Limited Systematicity Emerges: A Computational Cognitive Neuroscience Approach.

Conference Presentations

Schapiro, A.C., Trippe, A.M., Herd, S.A., O'Reilly, R.O., Rogers T.T., Norman, K.A. The Computational Mechanisms Underlying Learning during Sleep. Poster presented at the 13th Neural Computation and Psychology Workshop, San Sabastian, Spain, July, 2012

Herd, S., Friedman, N., Chatham, C., Hazy, T., Brant, A., O'Reilly, R. Towards Neural Network Models linking Genetics to Individual Differences in Executive Functions. Poster presented at the annual meeting of the Cognitive Neuroscience Society, SF, CA, March 2011

O'Reilly, R.C., Wyatte, D., Herd, S.A., Mingus, B., Kriete, T., Latimer, K. Bidirectional Biological Object Recognition and Figure-Ground Segmentation. Poster presented at Telluride Neuromorphic Engineering Conference, Telluride, CO, June 2011

Chatham, C.H., Brant, A.M., Herd, S.A., Hazy, TE, Miyake A, O'Reilly, R.C., Friedman N.P., From the Executive Network to Executive Functions – A PBWM Model of the n-back task. Talk at the annual meeting of the Computational Cognitive Neuroscience Society, SF, CA, March 2009

Sager, N., Herd, S.A. and Colunga, E. Modeling the development of bilingual and second language reading. Poster presented at the Cognitive Science Society Meeting, Nashville, TN, August 2007

Herd, S.A. & O'Reilly, R.C. A Neural Feature Integration Theory: Neural Network Models of Conjunctive Visual Search. Poster presented at the Society for Neuroscience Annual Meeting, San Diego, CA, November 2004

Herd, S.A. & O'Reilly, R.C.Visual Search: A Biologically Plausible Serial-Parallel Model. Poster Presented at the Computational Cognitive Neuroscience Meeting, Alicante, Spain, July 2003

Herd, S.A., Banich, M.T. & O'Reilly, R.C. Cognitive Control in the Stroop Task as Top-Down Bias: a computational model. Poster presented at the Cognitive Neuroscience Society Annual Meeting, San Francisco, April, 2002

Invited Talks

Computational Cognitive Neuroscience. Talk presented at University of Colorado Cognitive Science Club, 10/2012

Computational modeling of executive function. Talk series for the Determinants of Executive Function and Dysfunction center, University of Colorado, Spring 2009

Can science explain human consciousness? Talk presented at University of Colorado Cognitive Science Club, 10/4/2007

Serial-Parallel Models of Visual Search: Theory and Experiments. Talk presented at University of Colorado Department of Psychology, 5/03/2005

Neural Network Models of Visual Search. Talk presented at the Center for Neuroscience Supergroup, University of Colorado, 2/24/2004.

A Neural Network Model of Visual Search. Talk presented at University of Colorado Department of Psychology 4/14/2003

Accounting for fMRI data on the Stroop Task Using a Neural Network Model. Talk presented at Ekstrand Memorial Mini-Convention, Dept. of Psychology, University of Colorado, 4/29/2002

Neural Network Models of Cognitive Control on the Stroop Task. Talk presented at Ekstrand Memorial Mini-Convention, Dept. of Psychology, University of Colorado, 4/23/2001

Problem Solving in the Tower of Hanoi: Model and Experiment. Talk presented at Ekstrand Memorial Mini-Convention, Dept. of Psychology, University of Colorado, 4/24/2000

Teaching:

Instructor:
Computational Cognitive Neuroscience, Spring 2010, 2009, 2007
Introduction to Cognitive Psychology, Fall 2005

Lab Instructor:
Psychology of Learning, Dr. Jerry Rudy, Spring 2005
Computational Cognitive Neuroscience, Dr. Yuko Munakata, Fall 2003
Computational Cognitive Neuroscience, Dr. Randall O'Reilly, Spring 2001
Cognitive Psychology, Dr. Edward Crothers, Spring 2000

Teaching Assistant,
General Psychology, Dr. Brett King, Fall 1999

Lab Teaching Assistant:
Physics I: Mechanics (Calculus Track), Dr. John Howell, Fall 1995
Physics I: Mechanics (Calculus Track), Dr. John Howell, Fall 1994

Tutor,
Calculus I and II, 1994-1995

Research Experience:
Postdoctoral Professional Research Assistant, 2006-2012
Research Assistant, 2000-2005
Computational Cognitive Neuroscience laboratory
Department of Psychology and Neuroscience
University of Colorado at Boulder
Supervisor: Randy O'Reilly
Topics: Varied, including neural network models and theories of the following: Executive function, cognitive control, planning, vision, motor control, timing, audition, learning time sequences, error-driven learning, sensemaking, spatial processing, working memory, attention, reward prediction and dopamine function, and decision-making; and experimental design in visual search and effort/reward prediction for decision-making.

Research Assistant, 1996
Oak Ridge National Laboratories
Supervisor: Jeffrey Holmes
Topic: Computational simulation of climate change

Professional Affiliations:
Cognitive Neuroscience Society
Society for Neuroscience
Computational Neuroscience Society
American Psychological Association
Phi Beta Kappa

Awards and Honors:

University of Colorado Graduate Fellowship
Highest available College Honors, Earlham College, 1997
Highest available Departmental Honors, Physics, Earlham College, 1997
Outstanding Student in Physics Award, Earlham College, 1997
National Merit Scholar, 1993