Will human-like machines make human-like mistakes?

doi:10.1017/S0140525X1700019X, e270

Evan J. Livesey, Micah B. Goldwater, and Ben Colagiuri
School of Psychology, The University of Sydney, NSW 2006, Australia.
evan.livesey@sydney.edu.au micah.goldwater@sydney.edu.au
ben.colagiuri@sydney.edu.au
http://sydney.edu.au/science/people/evan.livesey.php
http://sydney.edu.au/science/people/micah.goldwater.php
http://sydney.edu.au/science/people/ben.colagiuri.php

Abstract: Although we agree with Lake et al.'s central argument, there are numerous flaws in the way people use causal models. Our models are often incorrect, resistant to correction, and applied inappropriately to new situations. These deficiencies are pervasive and have real-world consequences. Developers of machines with similar capacities should proceed with caution.

Lake et al. present a compelling case for why causal model-building is a key component of human learning, and we agree that beliefs about causal relations need to be captured by any convincingly human-like approach to artificial intelligence (AI). Knowledge of physical relations between objects and psychological relations between agents brings huge advantages. It provides a wealth of transferable information that allows humans to quickly apprehend a new situation. As such, combining the computational power of deep-neural networks with model-building capacities could indeed bring solutions to some of the world's most pressing problems. However, as advantageous as causal model-building might be, it also brings problems that can lead to flawed learning and reasoning. We therefore ask, would making machines "human-like" in their development of causal models also make those systems flawed in human-like ways?

Applying a causal model, especially one based on intuitive understanding, is essentially a gamble. Even though we often feel like we understand the physical and psychological relations surrounding us, our causal knowledge is almost always incomplete and sometimes completely wrong (Rozenblit & Keil 2002). These errors may be an inevitable part of the learning process by which models are updated based on experience. However, there are many examples in which incorrect causal models persist, despite strong counterevidence. Take the supposed link between immunisation and autism. Despite the science and the author of the original vaccine-autism connection being widely and publicly discredited, many continue to believe that immunisation increases the risk of autism and their refusal to immunise has decreased the population's immunity to preventable diseases (Larson et al. 2011; Silverman & Hendrix 2015).

Failures to revise false causal models are far from rare. In fact, they seem to be an inherent part of human reasoning. Lewandowsky and colleagues (2012) identify numerous factors that increase resistance to belief revision, including several that are societal-level (e.g., biased exposure to information) or motivational (e.g., vested interest in retaining a false belief). Notwithstanding the significance of these factors (machines too can be influenced by biases in data availability and the motives of their human developers), it is noteworthy that people still show resistance to updating their beliefs even when these sources of bias are removed, especially when new information conflicts with the existing causal model (Taylor & Ahn 2012).

Flawed causal models can also be based on confusions that are less easily traced to specific falsehoods. Well-educated adults regularly confuse basic ontological categories (Chi et al. 1994), distinctions between mental, biological, and physical phenomena that are fundamental to our models of the world and typically acquired in childhood (Carey 2011). A common example is the belief that physical energy possesses psychological desires and intentions – a belief that even some physics students appear to endorse (Svedholm & Lindeman 2013). These errors affect both our causal beliefs and our choices. Ontological confusions have been linked to people's acceptance of alternative medicine, potentially leading an individual to choose an ineffective treatment over evidence-based treatments, sometimes at extreme personal risk (Lindeman 2011).

Causal models, especially those that affect beliefs about treatment efficacy, can even influence physiological responses to medical treatments. In this case, known as the placebo effect, beliefs regarding a treatment can modulate the treatment response, positively or negatively, independently of whether a genuine treatment is delivered (Colagiuri et al. 2015). The placebo effect is caused by a combination of expectations driven by causal beliefs and associative learning mechanisms that are more analogous to the operations of simple neural networks. Associative learning algorithms, of the kind often used in neural networks, are surprisingly susceptible to illusory correlations, for example, when a treatment actually has no effect on a medical outcome (Matute et al. 2015). Successfully integrating two different mechanisms for knowledge generation (neural networks and causal models), when each individually may be prone to bias, is an interesting problem, not unlike the challenge of understanding the nature of human learning. Higher-level beliefs interact in numerous ways with basic learning and memory mechanisms, and the precise nature and consequences of these interactions remain unknown (Thorwart & Livesey 2016).

Even when humans hold an appropriate causal model, they often fail to use it. When facing a new problem, humans often erroneously draw upon models that share superficial properties with the current problem, rather than those that share key structural relations (Gick & Holyoak 1980). Even professional management consultants, whose job it is to use their prior experiences to help businesses solve novel problems, often fail to retrieve the most relevant prior experience to the new problem (Gentner et al. 2009). It is unclear whether an artificial system that possesses mental modelling capabilities would suffer the same limitations. On the one hand, they may be caused by human processing limitations. For example, effective model-based decision-making is associated with capacities for learning and transferring abstract rules (Don et al. 2016), and for cognitive control (Otto et al. 2015), which may potentially be far more powerful in future AI systems. On the other hand, the power of neural networks lies precisely in their ability to encode rich featural and contextual information. Given that experience with particular causal relations is likely to correlate with experience of more superficial features, a more powerful AI model generator may still suffer similar problems when faced with the difficult decision of which model to apply to a new situation.

Would human-like AI suffer human-like flaws, whereby recalcitrant causal models lead to persistence with poor solutions, or novel problems activate inappropriate causal models? Developers of AI systems should proceed with caution, as these properties of human causal modelling produce pervasive biases, and may be symptomatic of the use of mental models rather than the limitations on human cognition. Monitoring the degree to which AI systems show the same flaws as humans will be invaluable for shedding light on why human cognition is the way it is and, it is hoped, will offer some solutions to help us change our minds when we desperately need to.

Benefits of embodiment

doi:10.1017/S0140525X17000206, e271

Bruce James MacLennan
Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN 37996.
maclennan@utk.edu http://web.eecs.utk.edu/~mclennan

Abstract: Physical competence is acquired through animals' embodied interaction with their physical environments, and psychological

competence is acquired through situated interaction with other agents. The acquired neural models essential to these competencies are implicit and permit more fluent and nuanced behavior than explicit models. The challenge is to understand how such models are acquired and used to control behavior.

The target article argues for the importance of "developmental start-up software" (sects. 4.1 and 5.1), but neglects the nature of that software and how it is acquired. The embodied interaction of an organism with its environment, provides a foundation for its understanding of "intuitive physics" and physical causality. Animal nervous systems control their complex physical bodies in their complex physical environments in real time, and this competence is a consequence of innate developmental processes and, especially in more complex species, subsequent developmental processes that fine-tune neural control, such as prenatal and postnatal "motor babbling" (non-goal-directed motor activity) (Meltzoff & Moore 1997). Through these developmental processes, animals acquire a non-conceptual understanding of their bodies and physical environments, which provides a foundation for higher-order imaginative and conceptual physical understanding.

Animals acquire physical competence through interaction with their environments (both phylogenetic through evolution and ontogenetic through development), and robots can acquire physical competence similarly, for example, through motor babbling (Mahoor et al. 2016), and this is one goal of epigenetic and developmental robotics (Lungarella et al. 2003). In principle, comparable competence can be acquired by simulated physical agents behaving in simulated physical environments, but it is difficult to develop sufficiently accurate physical simulations so that agents acquire genuine physical competence (i.e., competence in the real world, not some simulated world). It should be possible to transfer physical competence from one agent to others that are sufficiently similar physically, but the tight coupling of body and nervous system suggests that physical competence will remain tied to a "form of life."

Animals are said to be *situated* because cognition primarily serves behavior, and behavior is always contextual. For most animals, situatedness involves interaction with other animals; it conditions the goals, motivations, and other factors that are causative in an animal's own behavior, and can be projected onto other agents, providing a foundation for "intuitive psychology." Psychological competence is grounded in the fact that animals are situated physical agents with interests, desires, goals, fears, and so on. Therefore, they have a basis for non-conceptual understanding of other agents (through imagination, mental simulation, projection, mirror neurons, etc.). In particular, they can project their experience of psychological causality onto other animals. This psychological competence is acquired through phylogenetic and ontogenetic adaptation.

The problem hindering AI systems from acquiring psychological competence is that most artificial agents do not have interests, desires, goals, fears, and so on that they can project onto others or use as a basis for mental simulation. For example, computer vision systems do not "care" in any significant way about the images they process. Because we can be injured and die, because we can feel fear and pain, we perceive immediately (i.e., without the mediation of conceptual thought) the significance of a man being dragged by a horse, or a family fleeing a disaster (Lake et al., Fig. 6). Certainly, through artificial evolution and reinforcement learning, we can train artificial agents to interact competently with other (real or simulated) agents, but because they are a different form of life, it will be difficult to give them the same cares and concerns as we have and that are relevant to many of our practical applications.

The target article does not directly address the important distinction between explicit and implicit models. *Explicit models* are the sort scientists construct, generally in terms of symbolic (lexical-level) variables; we expect to be able to understand explicit models conceptually, to communicate them in language, and to reason about them discursively (including mathematically). *Implicit models* are the sort that neural networks construct, generally in terms of large numbers of sub-symbolic variables, densely interrelated. Implicit models often allow an approximate emergent symbolic description, but such descriptions typically capture only the largest effects and interrelationships implicit in the sub-symbolic model. Therefore, they may lack the subtlety and context sensitivity of implicit models, which is why it is difficult, if not impossible, to capture expert behavior in explicit rules (Dreyfus & Dreyfus 1986). Therefore, terms such as "intuitive physics," "intuitive psychology," and "theory of mind" are misleading because they connote explicit models, but implicit models (especially those acquired by virtue of embodiment and situatedness) are more likely to be relevant to the sorts of learning discussed in the target article. It is less misleading to refer to *competencies*, because humans and other animals can use their physical and psychological understanding to behave competently even in the absence of explicit models.

The target article shows the importance of hierarchical compositionality to the physical competence of humans and other animals (sect. 4.2.1); therefore, it is essential to understand how hierarchical structure is represented in implicit models. Recognizing the centrality of embodiment can help, for our bodies are hierarchically articulated and our physical environments are hierarchically structured. The motor affordances of our bodies provide a basis for non-conceptual understanding of the hierarchical structure of objects and actions. However, it iss important to recognize that hierarchical decompositions need not be unique; they may be context dependent and subject to needs and interests, and a holistic behavior may admit multiple incompatible decompositions.

The target article points to the importance of simulation-based and imagistic inference (sect. 4.1.1). Therefore, we need to understand how they are implemented through implicit models. Fortunately, neural representations, such as topographic maps, permit analog transformations, which are better than symbolic digital computation for simulation-based and imagistic inference. The fact of neural implementation can reveal modes of information processing and control beyond the symbolic paradigm.

Connectionism consciously abandoned the explicit models of symbolic AI and cognitive science in favor of implicit, neural network models, which had a liberating effect on cognitive modeling, AI, and robotics. With 20-20 hindsight, we know that many of the successes of connectionism could have been achieved through existing statistical methods (e.g., Bayesian inference), without any reference to the brain, but they were not. Progress had been retarded by the desire for explicit, human-interpretable models, which connectionism abandoned in favor of neural plausibility. We are ill advised to ignore the brain again.

Understand the cogs to understand cognition

doi:10.1017/S0140525X17000218, e272

Adam H. Marblestone,[a] Greg Wayne,[b] and Konrad P. Kording[c]

[a]*Synthetic Neurobiology Group, MIT Media Lab, Cambridge, MA 02474;* [b]*DeepMind, London N1 9DR, UK;* [c]*Departments of Bioengineering and Neuroscience, University of Pennsylvania, Philadelphia, PA 19104*

adam.h.marblestone@gmail.com **gregwayne@gmail.com**
kording@upenn.edu **http://www.adammarblestone.org/**
www.kordinglab.com

Abstract: Lake et al. suggest that current AI systems lack the inductive biases that enable human learning. However, Lake et al.'s proposed biases may not directly map onto mechanisms in the developing brain. A convergence of fields may soon create a correspondence between biological neural circuits and optimization in structured architectures, allowing us to systematically dissect how brains learn.

References

[The letters "a" and "r" before author's initials stand for target article and response references, respectively]

Abelson, R. P. & Carroll, J. D. (1965) Computer simulation of individual belief systems. *The American behavioral scientist (pre-1986)* 8(9):24–30. [RJS]

Aitchison, L. & Lengyel, M. (2016) The Hamiltonian brain: Efficient probabilistic inference with excitatory-inhibitory neural circuit dynamics. *PLoS Computational Biology* 12(12):e1005186. [NK]

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C. & Qin, Y. (2004) An integrated theory of the mind. *Psychological Review* 111:1036–60. [RJS]

Anderson, M. L. (2003) Embodied cognition: A field guide. *Artificial Intelligence* 149 (1):91–130. [GB]

Andrychowicz, M., Denil, M., Gomez, S., Hoffman, M. W., Pfau, D., Schaul, T., Shillingford, B. & de Freitas, N. (2016). Learning to learn by gradient descent by gradient descent. Presented at the 2016 Neural Information Processing Systems conference, Barcelona, Spain, December 5–10, 2016. In: *Advances in neural information processing systems 29 (NIPS 2016)*, ed. D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon & R. Garnett, pp. 3981–89). Neural Information Processing Systems. [aBML, MB]

Anselmi, F., Leibo, J. Z., Rosasco, L., Mutch, J., Tacchetti, A. & Poggio, T. (2016) Unsupervised learning of invariant representations. *Theoretical Computer Science* 633:112–21. [aBML]

Ansermin, E., Mostafaoui, G., Beausse, N. & Gaussier, P. (2016). Learning to synchronously imitate gestures using entrainment effect. In: *From Animals to Animats 14: Proceedings of the 14th International Conference on Simulation of Adaptive Behavior (SAB 2016), Aberystwyth, United Kingdom, August 23–26, 2016,* ed. Tuci E, Giagkos A, Wilson M, Hallam J, pp. 219–31. Springer. [LM]

Arbib, M. A. & Fellous, J. M. (2004) Emotions: From brain to robot. *Trends in Cognitive Science* 8(12):554–61. [KBC]

Arnold, T. & Scheutz, M. (2016) Against the moral Turing test: Accountable design and the moral reasoning of autonomous systems. *Ethics and Information Technology* 18(2):103–15. doi:10.1007/s10676-016-9389-x. [PMP]

Asada, M. (2015) Development of artificial empathy. *Neuroscience Research* 90:41–50. [KBC]

Asada, M., Hosoda, K., Kuniyoshi, Y., Ishiguro, H., Inui, T., Yoshikawa, Y. & Yoshida, C. (2009) Cognitive developmental robotics: A survey. *IEEE Transactions on Autonomous Mental Development* 1(1):12–34. [P-YO]

Aurelius, M. (1937) *Meditations*, transl. G. Long. P. F. Collier & Son. [rBML]

Bach, J. (2009) *Principles of synthetic intelligence. PSI: An architecture of motivated cognition*. Oxford University Press. [CDG]

Bahdanau, D., Cho, K. & Bengio, Y. (2015) Neural machine translation by jointly learning to align and translate. Presented at the International Conference on Learning Representations (ICLR), San Diego, CA, May 7–9, 2015. *arXiv preprint 1409.0473.* Available at: http://arxiv.org/abs/1409.0473v3. [arBML]

Baillargeon, R. (2004) Infants' physical world. *Current Directions in Psychological Science* 13:89–94. [aBML]

Baillargeon, R., Li, J., Ng, W. & Yuan, S. (2009) An account of infants physical reasoning. In: *Learning and the infant mind,* ed. A. Woodward & A. Neeham, pp. 66–116. Oxford University Press. [aBML]

Baily, M. N. & Bosworth, B. P. (2014) US manufacturing: Understanding its past and its potential future. *The Journal of Economic Perspectives* 28(1):3–25. [DEM]

Baker, C. L., Jara-Ettinger, J., Saxe, R. & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour* 1:0064. [rBML]

Baker, C. L., Saxe, R. & Tenenbaum, J. B. (2009) Action understanding as inverse planning. *Cognition* 113(3):329–49. [arBML]

Baldassarre, G. (2011) What are intrinsic motivations? A biological perspective. In: *Proceedings of the International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob-2011)*, ed. A. Cangelosi, J. Triesch, I. Fasel, K. Rohlfing, F. Nori, P.-Y. Oudeyer, M. Schlesinger & Y. Nagai, pp. E1–8. IEEE. [GB]

Baldassarre, G., Caligiore, D. & Mannella, F. (2013a) The hierarchical organisation of cortical and basal-ganglia systems: A computationally-informed review and integrated hypothesis. In: *Computational and robotic models of the hierarchical organisation of behaviour*, ed. G. Baldassarre & M. Mirolli, pp. 237–70. Springer-Verlag. [GB]

Baldassarre, G., Mannella, F., Fiore, V. G., Redgrave, P., Gurney, K. & Mirolli, M. (2013b) Intrinsically motivated action-outcome learning and goal-based action recall: A system-level bio-constrained computational model. *Neural Networks* 41:168–87. [GB]

Baldassarre, G. & Mirolli, M., eds. (2013) *Intrinsically motivated learning in natural and artificial systems.* Springer. [GB, P-YO]

Baldassarre, G., Stafford, T., Mirolli, M., Redgrave, P., Ryan, R. M. & Barto, A. (2014) Intrinsic motivations and open-ended development in animals, humans, and robots: An overview. *Frontiers in Psychology* 5:985. [GB]

Baranes, A. & Oudeyer, P.-Y. (2013) Active learning of inverse models with intrinsically motivated goal exploration in robots. *Robotics and Autonomous Systems* 61(1):49–73. [P-YO]

Baranes, A. F., Oudeyer, P. Y. & Gottlieb, J. (2014) The effects of task difficulty, novelty and the size of the search space on intrinsically motivated exploration. *Frontiers in Neurosciences* 8:1–9. [P-YO]

Barros, P. & Wermter, S. (2016) Developing crossmodal expression recognition based on a deep neural model. *Adaptive Behavior* 24(5):373–96. [SW]

Barsalou, L. W. (1983) Ad hoc categories. *Memory & Cognition* 11(3):211–27. [aBML]

Barto, A. (2013) Intrinsic motivation and reinforcement learning. In: *Intrinsically motivated learning in natural and artificial systems*, ed. G. Baldassarre & M. Mirolli, pp. 17–47. Springer. [P-YO]

Bartunov, S. & Vetrov, D. P. (2016) Fast adaptation in generative models with generative matching networks. *arXiv preprint 1612.02192*. [SSH]

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P. & Friston, K. J. (2012) Canonical microcircuits for predictive coding. *Neuron* 76:695–711. http://doi.org/10.1016/j.neuron.2012.10.038. [aBML, DGe]

Bates, C. J., Yildirim, I., Tenenbaum, J. B. & Battaglia, P. W. (2015) Humans predict liquid dynamics using probabilistic simulation. In: *Proceedings of the 37th Annual Conference of the Cognitive Science Society, Pasadena, CA, July 22–25, 2015*, pp. 172–77. Cognitive Science Society. [aBML]

Battaglia, P., Pascanu, R., Lai, M. & Rezende, D. J. (2016) Interaction networks for learning about objects, relations and physics. Presented at the 2016 Neural Information Processing Systems conference, Barcelona, Spain, December 5–10, 2016. In: *Advances in neural information processing systems 29 (NIPS 2016)*, ed. D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon & R. Garnett, pp. 4502–10. Neural Information Processing Systems. [MB]

Battaglia, P. W., Hamrick, J. B. & Tenenbaum, J. B. (2013) Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences of the United States of America* 110(45):18327–32. [arBML, ED]

Baudiš, P. & Gailly, J.-l. (2012) PACHI: State of the art open source Go program. In: *Advances in computer games: 13th International Conference, ACG 2011, Tillburg, The Netherlands, November 20–22, 2011, Revised Selected Papers,* ed. H. Jaap van den Herik & A. Plast, pp. 24–38. Springer. [aBML]

Bauer, J., Dávila-Chacón, J. & Wermter, S. (2015) Modeling development of natural multi-sensory integration using neural self-organisation and probabilistic population codes. *Connection Science* 27(4):358–76. [SW]

Baxter, J. (2000) A model of inductive bias learning. *Journal of Artificial Intelligence Research* 12:149–98. [aBML]

Bayer, H. M. & Glimcher, P. W. (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47:129–41. [aBML]

Bellemare, M., Srinivasan, S., Ostrovski, G., Schaul, T., Saxton, D. & Munos, R. (2016) Unifying count-based exploration and intrinsic motivation. Presented at the 2016 Neural Information Processing Systems conference, Barcelona, Spain, December 5–10, 2016. In: *Advances in neural information processing systems 29 (NIPS 2016)*, ed. D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon & R. Garnett, pp. 1471–79. Neural Information Processing Systems. [MB, P-YO]

Bellemare, M. G., Naddaf, Y., Veness, J. & Bowling, M. (2013) The arcade learning environment: An evaluation platform for general agents. *Journal of Artificial Intelligence Research* 47:253–79. [aBML]

Bengio, J. (2009) Learning deep architectures for AI. *Foundations and Trends in Machine Learning* 2(1):1–127. [MBu]

Bengio, Y. (2016) Machines who learn. *Scientific American* 314(6):46–51. [KBC]

Bennis, W. M., Medin, D. L. & Bartels, D. M. (2010) The costs and benefits of calculation and moral rules. *Perspectives on Psychological Science* 5(2):187–202. doi:10.1177/1745691610362354. [PMP]

Berdahl, C. H. (2010) A neural network model of Borderline Personality Disorder. *Neural Networks* 23(2):177–88. [KBC]

Berlyne, D. E. (1966) Curiosity and exploration. *Science* 153(3731):25–33. doi:10.1126/science.153.3731.25 [aBML, CDG]

Berthiaume, V. G., Shultz, T. R. & Onishi, K. H. (2013) A constructivist connectionist model of transitions on false-belief tasks. *Cognition* 126(3): 441–58. [aBML]

Berwick, R. C. & Chomsky, N. (2016) *Why only us: Language and evolution*. MIT Press. [aBML]

Bever, T. G. & Poeppel, D. (2010) Analysis by synthesis: A (re-) emerging program of research for language and vision. *Biolinguistics* 4:174–200. [aBML]

Bi, G.-Q. & Poo, M.-M. (2001) Synaptic modification by correlated activity: Hebb's postulate revisited. *Annual Review of Neuroscience* 24:139–66. [aBML]

Biederman, I. (1987) Recognition-by-components: A theory of human image understanding. *Psychological Review* 94(2):115–47. [aBML]

Bienenstock, E., Cooper, L. N. & Munro, P. W. (1982) Theory for the development of neuron selectivity: Orientation specificity and binocular interaction in visual cortex. *The Journal of Neuroscience* 2(1):32–48. [aBML]

Bienenstock, E., Geman, S. & Potter, D. (1997) Compositionality, MDL priors, and object recognition. Presented at the 1996 Neural Information Processing Systems conference, Denver, CO, December 2–5, 1996. In: *Advances in neural* *information processing systems 9*, ed. M. C. Mozer, M. I. Jordan & T. Petsche, pp. 838–44. Neural Information Processing Systems Foundation. [aBML]

Blackburn, S. (1984) *Spreading the word: Groundings in the philosophy of language*. Oxford University Press. [NC]

Block, N. (1978) Troubles with functionalism. *Minnesota Studies in the Philosophy of Science* 9:261–325. [LRC]

Bloom, P. (2000) *How children learn the meanings of words*. MIT Press. [aBML]

Blumberg, M. S. (2005) *Basic instinct: The genesis of behavior.* Basic Books. [AHM]

Blundell, C., Uria, B., Pritzel, A., Li, Y., Ruderman, A., Leibo, J. Z., Rae, J., Wierstra, D. & Hassabis, D. (2016) Model-free episodic control. *arXiv preprint 1606.04460*. Available at: https://arxiv.org/abs/1606.04460. [aBML, MB]

Bobrow, D. G. & Winograd, T. (1977) An overview of KRL, a knowledge representation language. *Cognitive Science* 1:3–46. [aBML]

Boden, M. A. (1998) Creativity and artificial intelligence. *Artificial Intelligence* 103:347–56. [aBML]

Boden, M. A. (2006) *Mind as machine: A history of cognitive science*. Oxford University Press. [aBML]

Bonawitz, E., Denison, S., Griffiths, T. L. & Gopnik, A. (2014) Probabilistic models, learning algorithms, and response variability: Sampling in cognitive development. *Trends in Cognitive Sciences* 18:497–500. [aBML]

Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E. & Schulz, L. (2011) The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition* 120(3):322–30. Available at: http://doi.org/10.1016/j.cognition.2010.10.001. [MHT]

Bostrom, N. (2014) *Superintelligence: Paths, dangers, strategies*. Oxford University Press. ISBN 978-0199678112. [KBC]

Bottou, L. (2014) From machine learning to machine reasoning. *Machine Learning* 94(2):133–49. [aBML]

Botvinick, M. M. & Cohen, J. D. (2014) The computational and neural basis of cognitive control: Charted territory and new frontiers. *Cognitive Science* 38:1249–85. [MB]

Botvinick, M., Weinstein, A., Solway, A. & Barto, A. (2015) Reinforcement learning, efficient coding, and the statistics of natural tasks. *Current Opinion in Behavioral Sciences* 5:71–77. [MB]

Bouton, M. E. (2004) Context and behavioral processes in extinction. *Learning & Memory* 11:485–94. [aBML]

Boyd, R., Richerson, P. J. & Henrich, J. (2011) The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences of the United States of America* 108(suppl 2):10918–25. [MHT]

Braud, R., Mostafaoui, G., Karaouzene, A. & Gaussier, P. (2014). Simulating the emergence of early physical and social interactions: A developmental route through low level visuomotor learning. In: *From Animal to Animats 13: Proceedings of the 13th International Conference on Simulation of Adaptive Behavior, Castellon, Spain, July 2014*, ed. A. P. del Pobil, E. Chinalleto, E. Martinez-Martin, J. Hallam, E. Cervera & A. Morales, pp. 154–65. Springer. [LM]

Breazeal, C. & Scassellati, B. (2002). Robots that imitate humans. *Trends in Cognitive Sciences* 6(11):481–87. [LM]

Brenner, L. (2016) Exploring the psychosocial impact of Ekso Bionics Technology. *Archives of Physical Medicine and Rehabilitation* 97(10):e113. [DEM]

Briegel, H. J. (2012) On creative machines and the physical origins of freedom. *Scientific Reports* 2:522. [KBC]

Briggs, F. & Usrey, W. M. (2007) A fast, reciprocal pathway between the lateral geniculate nucleus and visual cortex in the macaque monkey. *The Journal of Neuroscience* 27(20):5431–36. [DG]

Buchsbaum, D., Gopnik, A., Griffiths, T. L. & Shafto, P. (2011) Children's imitation of causal action sequences is influenced by statistical and pedagogical evidence. *Cognition* 120(3):331–40. Available at: http://doi.org/10.1016/j.cognition.2010.12.001. [MHT]

Buckingham, D. & Shultz, T. R. (2000) The developmental course of distance, time, and velocity concepts: A generative connectionist model. *Journal of Cognition and Development* 1(3):305–45. [aBML]

Buesing, L., Bill, J., Nessler, B. & Maass, W. (2011) Neural dynamics as sampling: A model for stochastic computation in recurrent networks of spiking neurons. *PLoS Computational Biology* 7:e1002211. [aBML]

Burrell, J. (2016) How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society* 3(1):1–12. doi:10.1177/2053951715622512. [PMP]

Buscema, M. (1995) Self-reflexive networks: Theory – topology – Applications. *Quality and Quantity* 29(4):339–403. [MBu]

Buscema, M. (1998) Metanet*: The theory of independent judges. *Substance Use and Misuse* 32(2):439–61. [MBu]

Buscema, M. (2013) Artificial adaptive system for parallel querying of multiple databases. In: *Intelligent data mining in law enforcement analytics*, ed. M. Buscema & W. J. Tastle, pp. 481–511. Springer. [MBu]

Buscema, M., Grossi, E., Montanini, L. & Street, M. E. (2015) Data mining of determinants of intrauterine growth retardation revisited using novel algorithms generating semantic maps and prototypical discriminating variable profiles. *PLoS One* 10(7):e0126020. [MBu]

Buscema, M., Tastle, W. J. & Terzi, S. (2013) Meta net: A new meta-classifier family. In: *Data mining applications using artificial adaptive systems*, ed. W. J. Tastle, pp. 141–82. Springer. [MBu]

Buscema, M., Terzi, S. & Tastle, W.J. (2010). A new meta-classifier. In: *2010 Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS), Toronto, ON, Canada*, pp. 1–7. IEEE. [MBu]

Bushdid, C., Magnasco, M. O., Vosshall, L. B. & Keller, A. (2014) Humans can discriminate more than 1 trillion olfactory stimuli. *Science* 343(6177):1370–72. [DEM]

Caglar, L. R. & Hanson, S. J. (2016) Deep learning and attentional bias in human category learning. Poster presented at the Neural Computation and Psychology Workshop on Contemporary Neural Networks, Philadelphia, PA, August 8–10, 2016. [LRC]

Caligiore, D., Borghi, A., Parisi, D. & Baldassarre, G. (2010) TRoPICALS: A computational embodied neuroscience model of compatibility effects. *Psychological Review* 117(4):1188–228. [GB]

Caligiore, D., Pezzulo, G., Baldassarre, G., Bostan, A. C., Strick, P. L., Doya, K., Helmich, R. C., Dirkx, M., Houk, J., Jörntell, H., Lago-Rodriguez, A., Galea, J. M., Miall, R. C., Popa, T., Kishore, A., Verschure, P. F. M. J., Zucca, R. & Herreros, I. (2016) Consensus paper: Towards a systems-level view of cerebellar function: The interplay between cerebellum, basal ganglia, and cortex. *The Cerebellum* 16(1):203–29. doi: 10.1007/s12311-016-0763-3. [GB]

Calimera, A., Macii, E. & Poncino, M. (2013) The human brain project and neuromorphic computing. *Functional Neurology* 28(3):191–96. [KBC]

Cangelosi, A. & Schlesinger, M. (2015) *Developmental robotics: From babies to robots*. MIT Press. [P-YO, SW]

Cardon, A. (2006) Artificial consciousness, artificial emotions, and autonomous robots. *Cognitive Processes* 7(4):245–67. [KBC]

Carey, S. (1978) The child as word learner. In: *Linguistic theory and psychological reality*, ed. J. Bresnan, G. Miller & M. Halle, pp. 264–93. MIT Press. [aBML]

Carey, S. (2004) Bootstrapping and the origin of concepts. *Daedalus* 133(1):59–68. [aBML]

Carey, S. (2009) *The origin of concepts*. Oxford University Press. [arBML, KDF]

Carey, S. (2011) The origin of concepts: A précis. *Behavioral and Brain Sciences* 34 (03):113–62. [EJL]

Carey, S. & Bartlett, E. (1978) Acquiring a single new word. *Papers and Reports on Child Language Development* 15:17–29. [aBML]

Chavajay, P. & Rogoff, B. (1999) Cultural variation in management of attention by children and their caregivers. *Developmental Psychology* 35(4):1079. [JMC]

Chen, X. & Yuille, A. L. (2014) Articulated pose estimation by a graphical model with image dependent pairwise relations. In: *Advances in neural information processing systems 27 (NIPS 2014),* ed. Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence & K. Q. Weinberger, pp. 1736–44. Neural Information Processing Systems Foundation. [rBML]

Chen, Z. & Klahr, D. (1999) All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development* 70(5):1098–120. [KDF]

Chernova, S. & Thomaz, A. L. (2014) *Robot learning from human teachers.* Synthesis lectures on artificial intelligence and machine learning. Morgan & Claypool. [P-YO]

Chi, M. T., Slotta, J. D. & De Leeuw, N. (1994) From things to processes: A theory of conceptual change for learning science concepts. *Learning and Instruction* 4 (1):27–43. [EJL]

Chiandetti, C., Spelke, E. S. & Vallortigara, G. (2014) Inexperienced newborn chicks use geometry to spontaneously reorient to an artificial social partner. *Developmental Science* 18(6):972–78. doi:10.1111/desc.12277. [ESS]

Chouard, T. (2016) The Go files: AI computer wraps up 4–1 victory against human champion. (Online; posted March 15, 2016.) [aBML]

Christiansen, M. H. & Chater, N. (2016) *Creating language: Integrating evolution, acquisition, and processing.* MIT Press. [NC, SW]

Churchland, M. M., Cunningham, J. P., Kaufman, M. T., Foster, J. D., Nuyujukian, P., Ryu, S. I. & Shenoy, K. V. (2012) Neural population dynamics during reaching. *Nature* 487:51–56. [GB]

Ciresan, D., Meier, U. & Schmidhuber, J. (2012) Multi-column deep neural networks for image classification. In: *2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Providence, RI, June 16–21, 2012,* pp. 3642–49. IEEE. [aBML]

Clark, K. B. (2012) A statistical mechanics definition of insight. In: *Computational intelligence*, ed. A. G. Floares, pp. 139–62. Nova Science. ISBN 978-1-62081-901-2. [KBC]

Clark, K. B. (2014) Basis for a neuronal version of Grover's quantum algorithm. *Frontiers in Molecular Neuroscience* 7:29. [KBC]

Clark, K. B. (2015) Insight and analysis problem solving in microbes to machines. *Progress in Biophysics and Molecular Biology* 119:183–93. [KBC]

Clark, K. B. (in press-a) Classical and quantum Hebbian learning in modeled cognitive processing. *Frontiers in Psychology*. [KBC]

Clark, K. B. (in press-b) Neural field continuum limits and the partitioning of cognitive-emotional brain networks. *Molecular and Cellular Neuroscience*. [KBC]

Clark, K. B. (in press-c) Psychometric "Turing test" of general intelligences in social robots. *Information Sciences*. [KBC]

Clark, K. B. & Hassert, D. L. (2013) Undecidability and opacity of metacognition in animals and humans. *Frontiers in Psychology* 4:171. [KBC]

Cleeremans, A. (1993) *Mechanisms of implicit learning: Connectionist models of sequence processing*. MIT Press. [LRC]

Clegg, J. M., Wen, N. J. & Legare, C. H. (2017) Is non-conformity WEIRD? Cultural variation in adults' beliefs about children's competency and conformity. *Journal of Experimental Psychology: General* 146(3):428–41. [JMC]

Cohen, E. H. & Tong, F. (2015) Neural mechanisms of object-based attention. *Cerebral Cortex* 25(4):1080–92. http://doi.org/10.1093/cercor/bht303. [DGe]

Colagiuri, B., Schenk, L. A., Kessler, M. D., Dorsey, S. G. & Colloca, L. (2015) The placebo effect: from concepts to genes. *Neuroscience* 307:171–90. [EJL]

Colby, K. M. (1975) *Artificial paranoia: Computer simulation of paranoid processes.* Pergamon. [RJS]

Collins, A. G. E. & Frank, M. J. (2013) Cognitive control over learning: Creating, clustering, and generalizing task-set structure. *Psychological Review* 120 (1):190–229. [aBML]

Collins, S., Ruina, A., Tedrake, R. & Wisse, M. (2005) Efficient bipedal robots based on passive-dynamic walkers. *Science* 307(5712):1082–85. [P-YO]

Cook, C., Goodman, N. D. & Schulz, L. E. (2011) Where science starts: Spontaneous experiments in preschoolers' exploratory play. *Cognition* 120(3):341–49. [arBML]

Cooper, R. P. (2016) Executive functions and the generation of "random" sequential responses: A computational account. *Journal of Mathematical Psychology* 73:153–68. doi: 10.1016/j.jmp.2016.06.002. [RPK]

Correa-Chávez, M. & Rogoff, B. (2009) Children's attention to interactions directed to others: Guatemalan Mayan and European American patterns. *Developmental Psychology* 45(3):630. [JMC]

Corriveau, K. H. & Harris, P. L. (2010) Preschoolers (sometimes) defer to the majority when making simple perceptual judgments. *Developmental Psychology* 26:437–45. [JMC]

Corriveau, K. H., Kim, E., Song, G. & Harris, P. L. (2013) Young children's deference to a consensus varies by culture and judgment setting. *Journal of Cognition and Culture* 13(3–4):367–81. [JMC, rBML]

Coutinho, E., Deng, J. & Schuller, B. (2014) Transfer learning emotion manifestation across music and speech. In: *Proceedings of the 2014 International Joint Conference on Neural Networks (IJCNN), Beijing, China.* pp. 3592–98. IEEE. [SW]

Crick, F. (1989) The recent excitement about neural networks. *Nature* 337:129–32. [aBML]

Csibra, G. (2008) Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition* 107:705–17. [aBML]

Csibra, G., Biro, S., Koos, O. & Gergely, G. (2003) One-year-old infants use teleological representations of actions productively. *Cognitive Science* 27:111–33. [aBML]

Csibra, G. & Gergely, G. (2009) Natural pedagogy. *Trends in Cognitive Sciences* 13 (4):148–53. [MHT]

Dalrymple, D. (2016) *Differentiable programming*. Available at: https://www.edge.org/response-detail/26794. [aBML]

Davies, J. (2016) Program good ethics into artificial intelligence. *Nature* 538(7625). Available at: http://www.nature.com/news/program-good-ethics-into-artificial-intelligence-1.20821. [KBC]

Davis, E. & Marcus, G. (2014) The scope and limits of simulation in cognition. *arXiv preprint 1506.04956*. Available at: arXiv: http://arxiv.org/abs/1506.04956. [ED]

Davis, E. & Marcus, G. (2015) Commonsense reasoning and commonsense knowledge in artificial Intelligence. *Communications of the ACM* 58(9):92–103. [aBML]

Davis, E. & Marcus, G. (2016) The scope and limits of simulation in automated reasoning. *Artificial Intelligence* 233:60–72. [ED]

Davoodi, T., Corriveau, K. H. & Harris, P. L. (2016) Distinguishing between realistic and fantastical figures in Iran. *Developmental Psychology* 52(2):221. [JMC, rBML]

Daw, N. D., Niv, Y. & Dayan, P. (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience* 8(12):1704–11. doi:10.1038/nn1560. [aBML, RPK]

Day, S. B. & Gentner, D. (2007) Nonintentional analogical inference in text comprehension. *Memory & Cognition* 35:39–49. [KDF]

Dayan, P., Hinton, G. E., Neal, R. M. & Zemel, R. S. (1995) The Helmholtz machine. *Neural Computation* 7(5):889–904. [aBML]

Deacon, T. (2012) *Incomplete nature: How mind emerged from matter*. W.W. Norton. [DCD]

Deacon, T. W. (1998) *The symbolic species: The co-evolution of language and the brain*. WW Norton. [aBML]

Dehghani, M., Tomai, E., Forbus, K. & Klenk, M. (2008) An integrated reasoning approach to moral decision-making. In: *Proceedings of the 23rd AAAI National Conference on Artificial Intelligence, vol. 3*, pp. 1280–86. AAAI Press. [KDF]

DeJong, G. & Mooney, R. (1986) Explanation-based learning: An alternative view. *Machine Learning* 1(2):145–76. [LRC]

Denil, M., Agrawal, P., Kulkarni, T. D., Erez, T., Battaglia, P. & de Freitas, N. (2016). Learning to perform physics experiments via deep reinforcement learning. *arXiv preprint:1611.01843*. Available at: https://arxiv.org/abs/1611.01843. [MB]

Dennett, D. C. (1987) *The intentional stance*. MIT Press. [JMC]

Dennett, D. C. (2013) Aching voids and making voids [Review of the book *Incomplete nature: How mind emerged from matter* by T. Deacon]. *The Quarterly Review of Biology* 88(4):321–24. [DCD]

Dennett, D. C. (2017) *From bacteria to Bach and back: The evolution of minds*. W.W. Norton. [DCD]

Denton, E., Chintala, S., Szlam, A. & Fergus, R. (2015) Deep generative image models using a Laplacian pyramid of adversarial networks. Presented at the 2015 Neural Information Processing Systems conference, Montreal, QC, Canada, In: *Advances in neural information processing systems 28 (NIPS 2015)*, ed. C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama & R. Garnett [poster]. Neural Information Processing Systems Foundation. [aBML]

Di, G. Q. & Wu, S. X. (2015) Emotion recognition from sound stimuli based on back-projection neural networks and electroencephalograms. *Journal of the Acoustics Society of America* 138(2):994–1002. [KBC]

DiCarlo, J. J., Zoccolan, D. & Rust, N. C. (2012) How does the brain solve visual object recognition? *Neuron* 73(3):415–34. [NK]

Dick, P. K. (1968) *Do androids dream of electric sheep?* Del Ray-Ballantine. [DEM]

Dietvorst, B. J., Simmons, J. P. & Massey, C. (2015) Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General* 144(1):114–26. [DCD]

Dietvorst, B. J., Simmons, J. P. & Massey, C. (2016) Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2616787. [DCD]

Diuk, C., Cohen, A. & Littman, M. L. (2008) An object-oriented representation for efficient reinforcement learning. In: *Proceedings of the 25th International Conference on Machine Learning (ICML'08)*, Helsinki, Finland, pp. 240–47. ACM. [aBML]

DiYanni, C. J., Corriveau, K. H., Kurkul, K., Nasrini, J. & Nini, D. (2015) The role of consensus and culture in children's imitation of questionable actions. *Journal of Experimental Child Psychology* 137:99–110. [JMC]

Doeller, C. F., Barry, C. & Burgess, N. (2010) Evidence for grid cells in a human memory network. *Nature* 463(7281):657–61. doi:10.1038/nature08704. [ESS]

Doeller, C. F. & Burgess, N. (2008) Distinct error-correcting and incidental learning of location relative to landmarks and boundaries. *Proceedings of the National Academy of Sciences of the United States of America* 105(15):5909–14. [ESS]

Doeller, C. F., King, J. A. & Burgess, N. (2008) Parallel striatal and hippocampal systems for landmarks and boundaries in spatial memory. *Proceedings of the National Academy of Sciences of the United States of America* 105(15):5915–20. doi:10.1073/pnas.0801489105. [ESS]

Dolan, R. J. & Dayan, P. (2013) Goals and habits in the brain. *Neuron* 80:312–25. [aBML]

Don, H. J., Goldwater, M. B., Otto, A. R. & Livesey, E. J. (2016) Rule abstraction, model-based choice, and cognitive reflection. *Psychonomic Bulletin & Review* 23(5):1615–23. [EJL]

Donahue, J., Hendricks, L. A., Guadarrama, S., Rohrbach, M., Venugopalan, S., Saenko, K. & Darrell, T. (2015) Long-term recurrent convolutional networks for visual recognition and description. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, June 7-12, 2015,* pp. 2625–34. IEEE. [SW]

Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E. & Darrell, T. (2014) DeCAF: A deep convolutional activation feature for generic visual recognition. Presented at the International Conference on Machine Learning, Beijing, China, June 22–24, 2014. *Proceedings of Machine Learning Research* 32(1):647–55. [aBML]

Dörner, D. (2001) *Bauplan für eine Seele [Blueprint for a soul]*. Rowolt. [CDG]

Dörner, D. & Güss, C. D. (2013) PSI: A computational architecture of cognition, motivation, and emotion. *Review of General Psycholog* 17:297–317. doi:10.1037/a0032947. [CDG]

Doshi-Velez, F. & Kim, B. (2017) A roadmap for a rigorous science of interpretability. *arXiv preprint 1702.08608*. Available at: https://arxiv.org/abs/1702.08608. [rBML]

Doya, K. (1999) What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Networks* 12(7–8):961–74. [GB]

Dreyfus, H. & Dreyfus, S. (1986) *Mind over machine*. Macmillan. [BJM]

Duan, Y., Schulman, J., Chen, X., Bartlett, P. L., Sutskever, I. & Abbeel, P. (2016) RL^2: Fast reinforcement learning via slow reinforcement learning. *arXiv preprint 1611.02779*. Available at: https://arxiv.org/pdf/1703.07326.pdf. [MB]

Dunbar, K. (1995) How scientists really reason: Scientific reasoning in real-world laboratories. In: *The nature of insight*, ed. R. J. Sternberg & J. E. Davidson, pp. 365–95. MIT Press. [KDF]

Economides, M., Kurth-Nelson, Z., Lübbert, A., Guitart-Masip, M. & Dolan, R. J. (2015) Model-based reasoning in humans becomes automatic with training. *PLoS Computation Biology* 11:e1004463. [aBML]

Edelman, S. (2015) The minority report: Some common assumptions to reconsider in the modelling of the brain and behaviour. *Journal of Experimental & Theoretical Artificial Intelligence* 28(4):751–76. [aBML]

Eden, M. (1962) Handwriting and pattern recognition. *IRE Transactions on Information Theory* 8:160–66. [aBML]

Eickenberg, M., Gramfort, A., Varoquaux, G. & Thirion, B. (2016) Seeing it all: Convolutional network layers map the function of the human visual system. *NeuroImage* 2017;152:184–94. [NK]

Eliasmith, C., Stewart, T. C., Choo, X., Bekolay, T., DeWolf, T., Tang & Y. Rasmussen, D. (2012) A large-scale model of the functioning brain. *Science* 338(6111):1202–05. [aBML]

Eliasmith, C. & Trujillo, O. (2014) The use and abuse of large-scale brain models. *Current Opinion in Neurobiology* 25:1–6. [NK]

Elman, J. L. (1993) Learning and development in neural networks: The importance of starting small. *Cognition* 48(1):71–99. [SW]

Elman, J. L. (2005) Connectionist models of cognitive development: Where next? *Trends in Cognitive Sciences* 9(3):111–17. [aBML]

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D. & Plunkett, K. (1996) *Rethinking innateness*. MIT Press. [aBML]

Eslami, S. M., Heess, N., Weber, T., Tassa, Y., Kavukcuoglu, K. & Hinton, G. E. (2016) Attend, infer, repeat: Fast scene understanding with generative models. Presented at the 2016 Neural Information Processing Systems conference, Barcelona, Spain, December 5–10, 2016. In: *Advances in Neural Information Processing Systems 29 (NIPS 2016),* ed. D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon & R. Garnett, pp. 3225–33. Neural Information Processing Systems Foundation. [arBML, MB]

Eslami, S. M. A., Tarlow, D., Kohli, P. & Winn, J. (2014) Just-in-time learning for fast and flexible inference. Presented at the 2014 Neural Information Processing Systems conference, Montreal, QC, Canada, December 8–13, 2014. In: *Advances in neural information processing systems 27 (NIPS 2014),* ed. Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence & K. Q. Weinberger, pp. 1736–44. Neural Information Processing Systems Foundation. [aBML]

Fasolo, A. (2011) *The theory of evolution and its impact.* Springer. [DG]

Feigenbaum, E. & Feldman, J., eds. (1995) *Computers and thought.* AAAI Press. [RJS]

Flash, T., Hochner, B. (2005) Motor primitives in vertebrates and invertebrates. *Current Opinion in Neurobiology* 15(6):660–66. [P-YO]

Fodor, J. A. (1975) *The language of thought*. Harvard University Press. [aBML]

Fodor, J. A. (1981) *Representations: Philosophical essays on the foundations of cognitive science*. MIT Press. [LRC]

Fodor, J. A. & Pylyshyn, Z. W. (1988) Connectionism and cognitive architecture: A critical analysis. *Cognition* 28(1–2):3–71. [aBML, RPK, SSH]

Fogel, D. B. & Fogel, L. J. (1995) Evolution and computational intelligence. *IEEE Transactions on Neural Networks* 4:1938–41. [KBC]

Forbus, K. (2011) Qualitative modeling. *Wiley Interdisciplinary Reviews: Cognitive Science* 2(4):374–91. [KDF]

Forbus, K., Ferguson, R., Lovett, A. & Gentner, D. (2017) Extending SME to handle large-scale cognitive modeling. *Cognitive Science* 41(5):1152–201. doi:10.1111/cogs.12377. [KDF]

Forbus, K. & Gentner, D. 1997. Qualitative mental models: Simulations or memories? Presented at the Eleventh International Workshop on Qualitative Reasoning, Cortona, Italy, June 3–6, 1997. [KDF]

Forestier, S. & Oudeyer, P.-Y. (2016) Curiosity-driven development of tool use precursors: A computational model. In: *Proceedings of the 38th Annual Conference of the Cognitive Science Society, Philadelphia, PA,* ed. A. Papafragou, D. Grodner, D. Mirman & J. C. Trueswell, pp. 1859–1864. Cognitive Science Society. [P-YO]

Fornito, A., Zalesky, A. & Bullmore, E. (2016) *Fundamentals of brain network analysis.* Academic Press. [DG]

Fox, J., Cooper, R. P. & Glasspool, D. W. (2013) A canonical theory of dynamic decision-making. *Frontiers in Psychology* 4(150):1–19. doi: 10.3389/fpsyg.2013.00150. [RPK]

Frank, M. C. & Goodman, N. D. (2014) Inferring word meanings by assuming that speakers are informative. *Cognitive Psychology* 75:80–96. [MHT]

Frank, M. C., Goodman, N. D. & Tenenbaum, J. B. (2009) Using speakers' referential intentions to model early cross-situational word learning. *Psychological Science* 20:578–85. [aBML]

Franklin, S. (2007) A foundational architecture for artificial general intelligence. In: *Advances in artificial general intelligence: Concepts, architectures and algorithms: Proceedings of the AGI Workshop 2006*, ed. P. Want & B. Goertzel, pp. 36–54. IOS Press. [GB]

Freyd, J. (1983) Representing the dynamics of a static form. *Memory and Cognition* 11(4):342–46. [aBML]

Freyd, J. (1987) Dynamic mental representations. *Psychological Review* 94(4):427–38. [aBML]

Friedman, S. E. and Forbus, K. D. (2010) An integrated systems approach to explanation-based conceptual change. In: *Proceedings of the 24th AAAI Conference on Artificial Intelligence, Atlanta, GA, July 11–15, 2010*. AAAI Press. [KDF]

Fukushima, K. (1980) Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics* 36:193–202. [aBML]

Fung, P. (2015) Robots with heart. *Scientific American* 313(5):60–63. [KBC]

Funke, J. (2010) Complex problem solving: A case for complex cognition? *Cognitive Processing* 11:133–42. [CDG]

Gallese, V. & Lakoff, G. (2005) The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology* 22(3–4):455–79. [SW]

Gallistel, C. & Matzel, L. D. (2013) The neuroscience of learning: beyond the Hebbian synapse. *Annual Review of Psychology* 64:169–200. [aBML]

Gaussier, P., Moga, S., Quoy, M. & Banquet, J. P. (1998). From perception-action loops to imitation processes: A bottom-up approach of learning by imitation. *Applied Artificial Intelligence* 12(7–8):701–27. [LM]

Gazzaniga, M. (2004) *Cognitive neuroscience*. MIT Press. [MBu]

Gelly, S. & Silver, D. (2008) Achieving master level play in 9 × 9 computer Go. In: *Proceedings of the Twenty-third AAAI Conference on Artificial Intelligence, Chicago, Illinois, July 13–17, 2008,* pp. 1537–40. AAAI Press. [aBML]

Gelly, S. & Silver, D. (2011) Monte-Carlo tree search and rapid action value estimation in computer go. *Artificial Intelligence* 175(11):1856–75. [aBML]

Gelman, A., Carlin, J. B., Stern, H. S. & Rubin, D. B. (2004) *Bayesian data analysis*. Chapman & Hall/CRC. [aBML]

Gelman, A., Lee, D. & Guo, J. (2015) Stan: A probabilistic programming language for Bayesian inference and optimization. *Journal of Educational and Behavioral Statistics* 40:530–43. [aBML]

Gelman, S. A. (2009) Learning from others: Children's construction of concepts. *Annual Review of Psychology* 60:115–40. [MHT]

Geman, S., Bienenstock, E. & Doursat, R. (1992) Neural networks and the bias/variance dilemma. *Neural Computation* 4:1–58. [aBML]

Gentner, D. (1983) Structure-mapping: A theoretical framework for analogy. *Cognitive Science* 7:155–70. (Reprinted in A. Collins & E. E. Smith, eds. *Readings in cognitive science: A perspective from psychology and artificial intelligence*. Kaufmann.) [KDF]

Gentner, D. (2010) Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science* 34(5):752–75. [KDF]

Gentner, D., Loewenstein, J., Thompson, L. & Forbus, K. D. (2009) Reviving inert knowledge: Analogical abstraction supports relational retrieval of past events. *Cognitive Science* 33(8):1343–82. [EJL]

George, D. & Hawkins, J. (2009) Towards a mathematical theory of cortical microcircuits. *PLoS Computational Biology* 5(10):e1000532. Available at: http://doi.org/10.1371/journal.pcbi.1000532. [DGe]

Gershman, S. J. & Goodman, N. D. (2014) Amortized inference in probabilistic reasoning. In: *Proceedings of the 36th Annual Conference of the Cognitive Science Society, Quebec City, QC, Canada, July 23–26, 2014*, pp. 517–522. Cognitive Science Society. [aBML]

Gershman, S. J., Horvitz, E. J. & Tenenbaum, J. B. (2015) Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science* 34:273–78. [aBML]

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* 143:182–94. [aBML]

Gershman, S. J., Vul, E. & Tenenbaum, J. B. (2012) Multistability and perceptual inference. *Neural Computation* 24:1–24. [aBML]

Gerstenberg, T., Goodman, N. D., Lagnado, D. A. & Tenenbaum, J. B. (2015) How, whether, why: Causal judgments as counterfactual contrasts. In: *Proceedings of the 37th Annual Conference of the Cognitive Science Society, Pasadena, CA, July 22–25, 2015*, ed. D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings & P. P. Maglio, pp. 782–787. Cognitive Science Society. [aBML, ED]

Ghahramani, Z. (2015) Probabilistic machine learning and artificial intelligence. *Nature* 521:452–59. [aBML]

Giambene, G. (2005) *Queuing theory and telecommunications networks and applications.* Springer Science + Business Media. [DG]

Gibson, J. J. (1979) *The ecological approach to visual perception*. Houghton Mifflin. [DCD]

Gick, M. L. & Holyoak, K. J. (1980) Analogical problem solving. *Cognitive Psychology* 12(3):306–55. [EJL]

Gigerenzer, G. (2001) The adaptive toolbox. In: *Bounded rationality: The adaptive toolbox*, ed. G. Gigerenzer & R. Selten, pp. 37–50. MIT Press. [PMP]

Gigerenzer, G. & Gaissmaier, W. (2011) Heuristic decision making. *Annual Review of Psychology* 62:451–82. doi:10.1146/annurev-psych-120709-145346. [PMP]

Goldberg, A. E. (1995) *Constructions: A construction grammar approach to argument structure*. University of Chicago Press. [NC]

Gombrich, E. (1960) *Art and illusion*. Pantheon Books. [NC]

Goodfellow, I., Schlens, J. & Szegedy, C. (2015) Explaining and harnessing adversarial examples. Presented at International Conference on Learning Representations (ICLR), San Diego, CA, May 7–9, 2015. *arXiv preprint 1412.6572*. Available at: https://arxiv.org/abs/1412.6572. [KDF]

Goodman, N. D. & Frank, M. C. (2016) Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences* 20(11):818–29. [MHT]

Goodman, N. D., Mansinghka, V. K., Roy, D. M., Bonawitz, K. & Tenenbaum, J. B. (2008) Church: A language for generative models. In: *Proceedings of the Twenty-Fourth Annual Conference on Uncertainty in Artificial Intelligence, Helsinki, Finland, July 9–12, 2008*, pp. 220–29. AUAI Press. [aBML]

Goodman, N. D., Tenenbaum, J. B., Feldman, J. & Griffiths, T. L. (2008) A rational analysis of rule-based concept learning. *Cognitive Science* 32(1):108–54. [rBML]

Goodman, N. D., Tenenbaum, J. B. & Gerstenberg, T. (2015). Concepts in a probabilistic language of thought. In: *The conceptual mind: New directions in the study of concepts*, ed. E. Margolis & S. Laurence, pp. 623–54. MIT Press. [rBML]

Goodman, N. D., Ullman, T. D. & Tenenbaum, J. B. (2011) Learning a theory of causality. *Psychological Review* 118(1):110–19. [rBML]

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T. & Danks, D. (2004) A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review* 111(1):3–32. [arBML]

Gopnik, A. & Meltzoff, A. N. (1999) *Words, thoughts, and theories*. MIT Press. [aBML]

Gottlieb, J., Oudeyer, P-Y., Lopes, M. & Baranes, A. (2013) Information seeking, curiosity and attention:Computational and neural mechanisms. *Trends in Cognitive Science* 17(11):585–96. [P-YO]

Graham, D. J. (2014) Routing in the brain. *Frontiers in Computational Neuroscience* 8:44. [DG]

Graham, D. J. and Rockmore, D. N. (2011) The packet switching brain. *Journal of Cognitive Neuroscience* 23(2):267–76. [DG]

Granger, R. (2006) Engines of the brain: The computational instruction set of human cognition. *AI Magazine* 27(2):15. [DG]

Graves, A. (2014) Generating sequences with recurrent neural networks. *arXiv preprint 1308.0850*. Available at: http://arxiv.org/abs/1308.0850. [aBML]

Graves, A., Mohamed, A.-R. & Hinton, G. (2013) Speech recognition with deep recurrent neural networks. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Vancouver, BC, Canada, May 26–31, 2013*, pp. 6645–49. IEEE. [aBML]

Graves, A., Wayne, G. & Danihelka, I. (2014) Neural Turing machines. *arXiv preprint 1410.5401v1*. Available at: http://arxiv.org/abs/1410.5401v1. [arBML]

Graves, A., Wayne, G., Reynolds, M., Harley, T., Danihelka, I., Grabska-Barwińska, A., Colmenarejo, S. G., Grefenstette, E., Ramalho, T., Agapiou, J., Badia, A. P., Hermann, K. M., Zwols, Y., Ostrovski, G., Cain, A., King, H., Summerfield, C., Blunsom, P., Kayukcuoglu, K. & Hassabis, D. (2016) Hybrid computing using a neural network with dynamic external memory. *Nature* 538(7626):471–76. [arBML, MB]

Gray, H. M., Gray, K. & Wegner, D. M. (2007) Dimensions of mind perception. *Science* 315(5812):619. [rBML, SW]

Gray, K. & Wegner, D. M. (2012) Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition* 125(1):125–30. [rBML]

Graybiel, A. M. (2005) The basal ganglia: learning new tricks and loving it. *Current Opinion in Neurobiology* 15(6):638–44. [GB]

Grefenstette, E., Hermann, K. M., Suleyman, M. & Blunsom, P. (2015). Learning to transduce with unbounded memory. Presented at the 2015 Neural Information Processing Systems conference. In: *Advances in Neural Information Processing Systems 28*, ed. C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama & R. Garnett. Neural Information Processing Systems Foundation. [arBML]

Gregor, K., Besse, F., Rezende, D. J., Danihelka, I. & Wierstra, D. (2016) Towards conceptual compression. Presented at the 2016 Neural Information Processing Systems conference, Barcelona, Spain, December 5–10, 2016. In: *Advances in Neural Information Processing Systems 29 (NIPS 2016)*, ed. D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon & R. Garnett [poster]. Neural Information Processing Systems Foundation. [aBML]

Gregor, K., Danihelka, I., Graves, A., Rezende, D. J. & Wierstra, D. (2015) DRAW: A recurrent neural network for image generation. Presented at the 32nd Annual International Conference on Machine Learning (ICML'15), Lille, France, July 7–9, 2015. *Proceedings of Machine Learning Research* 37:1462–71. [aBML]

Griffiths, T. L., Chater, N., Kemp, C., Perfors, A. & Tenenbaum, J. B. (2010) Probabilistic models of cognition: Exploring representations and inductive biases. *Trends in Cognitive Sciences* 14(8):357–64. [arBML]

Griffiths, T. L. & Tenenbaum, J. B. (2005) Structure and strength in causal induction. *Cognitive Psychology* 51(4):334–84. [rBML]

Griffiths, T. L. & Tenenbaum, J. B. (2009) Theory-based causal induction. *Psychological Review* 116(4):661–716. [rBML]

Griffiths, T. L., Vul, E. & Sanborn, A. N. (2012) Bridging levels of analysis for probabilistic models of cognition. *Current Directions in Psychological Science* 21:263–68. [aBML]

Grossberg, S. (1976) Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors. *Biological Cybernetics* 23:121–34. [aBML]

Grosse, R., Salakhutdinov, R., Freeman, W. T. & Tenenbaum, J. B. (2012) Exploiting compositionality to explore a large space of model structures. In: Proceedings of the Twenty-Eighth Annual Conference on *Uncertainty in Artificial Intelligence , Catalina Island , CA*, ed. N. de Freitas & K. Murphy , pp. 306–15. AUAI Press. [aBML]

Güçlü, U. & van Gerven, M. A. J. (2015) Deep neural networks reveal a gradient in the complexity of neural representations across the ventral stream. *Journal of Neuroscience* 35(27):10005–14. [NK]

Guergiuev, J., Lillicrap, T. P. & Richards, B. A. (2016) Toward deep learning with segregated dendrites. *arXiv preprint 1610.00161*. Available at: http://arxiv.org/pdf/1610.00161.pdf. [AHM]

Gülçehre, Ç. & Bengio, Y. (2016) Knowledge matters: Importance of prior information for optimization. *Journal of Machine Learning Research* 17(8):1–32. [SSH]

Guo, X., Singh, S., Lee, H., Lewis, R. L. & Wang, X. (2014) Deep learning for real-time Atari game play using offline Monte-Carlo tree search planning. In: *Advances in neural information processing systems 27 (NIPS 2014),* ed. Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence & K. Q. Weinberger [poster]. Neural Information Processing Systems Foundation. [aBML]

Güss, C. D., Tuason, M. T. & Gerhard, C. (2010) Cross-national comparisons of complex problem-solving strategies in two microworlds. *Cognitive Science* 34:489–520. [CDG]

Gweon, H., Tenenbaum, J. B. & Schulz, L. E. (2010) Infants consider both the sample and the sampling process in inductive generalization. *Proceedings of the National Academy of Sciences of the United States of America* 107:9066–71. [aBML]

Hafenbrädl, S., Waeger, D., Marewski, J. N. & Gigerenzer, G. (2016) Applied decision making with fast-and-frugal heuristics. *Journal of Applied Research in Memory and Cognition* 5(2):215–31. doi:10.1016/j.jarmac.2016.04.011. [PMP]

Hall, E. T. (1966) *The hidden dimension*. Doubleday. [SW]

Halle, M. & Stevens, K. (1962) Speech recognition: A model and a program for research. *IRE Transactions on Information Theory* 8(2):155–59. [aBML]

Hamlin, K. J. (2013) Moral judgment and action in preverbal infants and toddlers: Evidence for an innate moral core. *Current Directions in Psychological Science* 22:186–93. [aBML]

Hamlin, K. J., Ullman, T., Tenenbaum, J., Goodman, N. D. & Baker, C. (2013) The mentalistic basis of core social cognition: Experiments in preverbal infants and a computational model. *Developmental Science* 16:209–26. [aBML]

Hamlin, K. J., Wynn, K. & Bloom, P. (2007) Social evaluation by preverbal infants. *Nature* 450:57–60. [aBML]

Hamlin, K. J., Wynn, K. & Bloom, P. (2010) Three-month-olds show a negativity bias in their social evaluations. *Developmental Science* 13:923–29. [aBML]

Hamper, B. (2008) *Rivethead: Tales from the assembly line*. Grand Central. [DEM]

Hamrick, J. B., Ballard, A. J., Pascanu, R., Vinyals, O., Heess, N. & Battaglia, P. W. (2017) Metacontrol for adaptive imagination-based optimization. In: *Proceedings of the 5th International Conference on Learning Representations (ICLR)*. [MB]

Han, M. J., Lin, C. H. & Song, K. T. (2013) Robotic emotional expression generation based on mood transition and personality model. *IEEE Transactions on Cybernetics* 43(4):1290–303. [KBC]

Hannun, A., Case, C., Casper, J., Catanzaro, B., Diamos, G., Elsen, E., Prenger, R., Satheesh, S., Shubho, S., Coates, A. & Ng, A. Y. (2014) Deep speech: Scaling up end-to-end speech recognition. *arXiv preprint 1412.5567.* Available at: https://arxiv.org/abs/1412.5567. [aBML]

Hanson, S. J., (1995) Some comments and variations on back-propagation. In: *The handbook of back-propagation*, ed. Y. Chauvin & D. Rummelhart, pp. 292–323. Erlbaum. [LRC]

Hanson, S. J. (2002) On the emergence of rules in neural networks. *Neural Computation* 14(9):2245–68. [LRC]

Hanson, S. J. & Burr, D. J., (1990) What connectionist models learn: Toward a theory of representation in connectionist networks. *Behavioral and Brain Sciences* 13:471–518. [LRC]

Hanson, S. J., Caglar, L. R. & Hanson, C. (under review) The deep history of deep learning. [LRC]

Harkness, S., Blom, M., Oliva, A., Moscardino, U., Zylicz, P. O., Bermudez, M. R. & Super, C. M. (2007) Teachers' ethnotheories of the 'ideal student' in five western cultures. *Comparative Education* 43(1):113–35. [JMC]

Harlow, H. F. (1949) The formation of learning sets. *Psychological Review* 56(1):51–65. [aBML]

Harlow, H. F. (1950) Learning and satiation of response in intrinsically motivated complex puzzle performance by monkeys. *Journal of Comparative and Physiological Psychology* 43:289–94. [aBML]

Harris, P. L. (2012) *Trusting what you're told: How children learn from others*. Belknap Press of Harvard University Press. [JMC]

Haslam, N. (2006) Dehumanization: An integrative review. *Personality and Social Psychology Review* 10(3):252–64. [rBML]

Hasnain, S.K., Mostafaoui, G. & Gaussier, P. (2012). A synchrony-based perspective for partner selection and attentional mechanism in human-robot interaction. *Paladyn, Journal of Behavioral Robotics* 3(3):156–71. [LM]

Hasnain, S. K., Mostafaoui, G., Salesse, R., Marin, L. & Gaussier, P. (2013) Intuitive human robot interaction based on unintentional synchrony: A psycho-experimental study. In: *Proceedings of the IEEE 3rd Joint Conference on Development* *and Learning and on Epigenetic Robotics, Osaka, Japan, August 2013*, pp. 1–7. Hal Archives-ouvertes. [LM]

Hauser, M. D., Chomsky, N. & Fitch, W. T. (2002) The faculty of language: what is it, who has it, and how did it evolve? *Science* 298:1569–79. [aBML]

Hayes, P. J. (1974) Some problems and non-problems in representation theory. In: *Proceedings of the 1st summer conference on artificial intelligence and simulation of behaviour*, pp. 63–79. IOS Press. [LRC]

Hayes-Roth, B. & Hayes-Roth, F. (1979) A cognitive model of planning. *Cognitive Science* 3:275–310. [aBML]

He, K., Zhang, X., Ren, S. & Sun, J. (2016) Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, June 27–30, 2016. pp. 770–78. IEEE. [aBML]

Hebb, D. O. (1949) *The organization of behavior*. Wiley. [aBML]

Heess, N., Tarlow, D. & Winn, J. (2013) Learning to pass expectation propagation messages. Presented at the 25th International Conference on Neural Information Processing Systems, Lake Tahoe, NV, December 3–6, 2012. In: *Advances in Neural Information Processing Systems 25 (NIPS 2012),* ed. F. Pereira, C. J. C. Burges, L. Bottou & K. Q. Weinberger, pp. 3219–27. Neural Information Processing Systems Foundation. [aBML]

Heinrich, S. (2016) *Natural language acquisition in recurrent neural architectures*. Ph.D. thesis, Universität Hamburg, DE. [SW]

Henrich, J. (2015) *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton University Press. [JMC]

Henrich, J., Heine, S. J. & Norenzayan, A. (2010) The weirdest people in the world? *Behavioral and Brain Sciences* 33(2–3):61–83. [JMC]

Herrmann, E., Call, J., Hernandez-Lloreda, M. V., Hare, B. & Tomasello, M. (2007) Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science* 317(5843):1360–66. [DCD]

Herrmann, E., Hernandez-Lloreda, M. V., Call, J., Hare, B. & Tomasello, M. (2010) The structure of individual differences in the cognitive abilities of children and chimpanzees. *Psychological Science* 21(1):102–10. [DCD]

Hertwig, R. & Herzog, S. M. (2009) Fast and frugal heuristics: Tools of social rationality. *Social Cognition* 27(5):661–98. doi:10.1521/soco.2009.27.5.661. [PMP]

Hespos, S. J. & Baillargeon, R. (2008) Young infants' actions reveal their developing knowledge of support variables: Converging evidence for violation-of-expectation findings. *Cognition* 107:304–16. [aBML]

Hespos, S. J., Ferry, A. L. & Rips, L. J. (2009) Five-month-old infants have different expectations for solids and liquids. *Psychological Science* 20(5):603–11. [aBML]

Hinrichs, T. & Forbus, K. (2011) Transfer learning through analogy in games. *AI Magazine* 32(1):72–83. [KDF]

Hinton, G. E. (2002) Training products of experts by minimizing contrastive divergence. *Neural Computation* 14(8):1771–800. [aBML]

Hinton, G. E., Dayan, P., Frey, B. J. & Neal, R. M. (1995) The "wake-sleep" algorithm for unsupervised neural networks. *Science* 268(5214):1158–61. [aBML]

Hinton, G. E., Deng, L., Yu, D., Dahl, G. E., Mohamed, A.-r., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T. & Kingsbury, B. (2012) Deep neural networks for acoustic modeling in speech recognition. *IEEE Signal Processing Magazine* 29:82–97. [aBML]

Hinton, G. E., Osindero, S. & Teh, Y. W. (2006) A fast learning algorithm for deep belief nets. *Neural Computation 18*:1527–54. [aBML]

Hiolle, A., Lewis, M. & Cañamero, L. (2014) Arousal regulation and affective adaptation to human responsiveness by a robot that explores and learns a novel environment. *Frontiers in Neurorobotics* 8:17. [KBC]

Ho, Y-C. & Pepyne, D. L. (2002) Simple explanation of the no-free-lunch theorem and its implications. *Journal of Optimization Theory and Applications* 115:549–70. [AHM]

Hochreiter, S. A., Younger, S. & Conwell, P. R. (2001) Learning to learn using gradient descent. In: *International Conference on Artificial Neural Network—ICANN 2001,* ed. G. Dorffner, H. Bischoff & K. Hornik, pp. 87–94. Springer. [MB]

Hoffman, D. D. (2000) *Visual intelligence: How we create what we see*. W. W. Norton. [NC]

Hoffman, D. D. & Richards, W. A. (1984) Parts of recognition. *Cognition* 18:65–96. [aBML]

Hoffman, M., Yoeli, E. & Nowak, M. A. (2015) Cooperate without looking: Why we care what people think and not just what they do. *Proceedings of the National Academy of Sciences of the United States of America* 112(6):1727–32. doi:10.1073/pnas.1417904112. [PMP]

Hofstadter, D. R. (1985) *Metamagical themas: Questing for the essence of mind and pattern*. Basic Books. [aBML]

Hofstadter, D. R. (2001) Epilogue: Analogy as the core of cognition. In: *The analogical mind: perspectives from cognitive science*, ed. D. Gentner, K. J. Holyoak & B. N. Kokinov, pp. 499–538. MIT Press. [NC]

Horgan, T. & J. Tienson, (1996) *Connectionism and the philosophy of psychology*. MIT Press. [LRC]

Horst, J. S. & Samuelson, L. K. (2008) Fast mapping but poor retention by 24-month-old infants. *Infancy* 13(2):128–57. [aBML]

Houk, J. C., Adams, J. L. & Barto, A. G. (1995) A model of how the basal ganglia generate and use neural signals that predict reinforcement. In: *Models of information processing in the basal ganglia*, ed. J. C. Houk, J. L. Davids & D. G. Beiser, pp. 249–70. MIT Press. [GB]

Huang, Y. & Rao, R. P. (2014) Neurons as Monte Carlo samplers: Bayesian? inference and learning in spiking networks Presented at the 2014 Neural Information Processing Systems conference, Montreal, QC, Canada. In: *Advances in neural information processing systems 27 (NIPS 2014),* ed. Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence & K. Q. Weinberger, pp. 1943–51. Neural Information Processing Systems Foundation. [aBML]

Hubel, D. H. & Wiesel, T. N. (1959) Receptive fields of single neurons in the cat's striate cortex. *Journal of Physiology* 124:574–91. [ESS]

Hummel, J. E. & Biederman, I. (1992) Dynamic binding in a neural network for shape recognition. *Psychological Review* 99(3):480–517. [aBML]

Hurley, M., Dennett, D. C. & Adams, R., (2011) *Inside jokes: Using humor to reverse-engineer the mind.* MIT Press. [DCD]

Hutson, M. (2017) In bots we distrust. *Boston Globe*, p. K4. [DCD]

Indiveri, G. & Liu, S.-C. (2015) Memory and information processing in neuromorphic systems. *Proceedings of the IEEE* 103(8):1379–97. [KBC]

Indurkhya, B. & Misztal-Radecka, J. (2016) Incorporating human dimension in autonomous decision-making on moral and ethical issues. In: *Proceedings of the AAAI Spring Symposium: Ethical and Moral Considerations in Non-human Agents, Palo Alto, CA,* ed. B. Indurkhya & G. Stojanov. AAAI Press. [PMP]

Irvine, A. D. & Deutsch, H. (2016) Russell's paradox. In: *The Stanford encyclopedia of philosophy* (Winter 2016 Edition), ed. E. N. Zalta. Available at: https://plato.stanford.edu/archives/win2016/entries/russell-paradox. [NC]

Jackendoff, R. (2003) *Foundations of language*. Oxford University Press. [aBML]

Jaderberg, M., Mnih, V., Czarnecki, W. M., Schaul, T., Leibo, J. Z., Silver, D. & Kavukcuoglu, K. (2016) Reinforcement learning with unsupervised auxiliary tasks. Presented at the 5th International Conference on Learning Representations, Palais des Congrès Neptune, Toulon, France, April 24–26, 2017. *arXiv preprint 1611.05397*. Available at: https://arxiv.org/abs/1611.05397. [P-YO]

Jain, A., Tompson, J., Andriluka, M., Taylor, G. W. & Bregler, C. (2014). Learning human pose estimation features with convolutional networks. Presented at the International Conference on Learning Representations (ICLR), Banff, Canada, April 14–16, 2014. *arXiv preprint 1312.7302*. Available at: https://arxiv.org/abs/1312.7302. [rBML]

Jara-Ettinger, J., Gweon, H., Schulz, L. E. & Tenenbaum, J. B. (2016) The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in Cognitive Sciences* 20(8):589–604. doi:10.1016/j.tics.2016.05.011. [PMP]

Jara-Ettinger, J., Gweon, H., Tenenbaum, J. B. & Schulz, L. E. (2015) Children's understanding of the costs and rewards underlying rational action. *Cognition* 140:14–23. [aBML]

Jern, A. & Kemp, C. (2013) A probabilistic account of exemplar and category generation. *Cognitive Psychology* 66(1):85–125. [aBML]

Jern, A. & Kemp, C. (2015) A decision network account of reasoning about other peoples choices. *Cognition* 142:12–38. [aBML]

Johnson, M., Schuster, M., Le, Q. V., Krikun, M., Wu, Y., Chen, Z. & Hughes, M. (2016) Google's multilingual neural machine translation system: Enabling zero-shot translation. *arXiv preprint 1611.04558*. Available at: https://arxiv.org/abs/1611.04558. [SSH]

Johnson, S. C., Slaughter, V. & Carey, S. (1998) Whose gaze will infants follow? The elicitation of gaze-following in 12-month-olds. *Developmental Science* 1:233–38. [aBML]

Jonge, M. de & Racine, R. J. (1985) The effects of repeated induction of long-term potentiation in the dentate gyrus. *Brain Research* 328:181–85. [aBML]

Juang, B. H. & Rabiner, L. R. (1990) Hidden Markov models for speech recognition. *Technometric* 33(3):251–72. [aBML]

Kahneman, D. (2011) *Thinking, fast and slow*. Macmillan. [MB]

Kahou, S. E., Pal, C., Bouthillier, X., Froumenty, P., Gülçehre, Ç., Memisevic, R., Vincent, P., Courville, A. & Bengio, Y. (2013) Combining modality specific deep neural networks for emotion recognition in video. In: *Proceedings of the 15th ACM International Conference on Multimodal Interaction, Koogee Beach, Sydney, Australia,* pp. 543–50. ACM. [rBML]

Kaipa, K. N., Bongard, J. C. & Meltzoff, A. N. (2010) Self discovery enables robot social cognition: Are you my teacher? *Neural Networks* 23(8–9):1113–24. [KBC]

Karpathy, A. & Fei-Fei, L. (2017) Deep visual-semantic alignments for generating image descriptions. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39(4):664–76. [aBML]

Kawato, M., Kuroda, S. & Schweighofer, N. (2011) Cerebellar supervised learning revisited: biophysical modeling and degrees-of-freedom control. *Current Opinion in Neurobiology* 21(5):791–800. [GB]

Keller, N. & Katsikopoulos, K. V. (2016) On the role of psychological heuristics in operational research; and a demonstration in military stability operations.

European Journal of Operational Research 249(3):1063–73. doi:10.1016/j.ejor.2015.07.023. [PMP]

Kellman, P. J. & Spelke, E. S. (1983) Perception of partly occluded objects in infancy. *Cognitive Psychology* 15(4):483–524. [rBML]

Kemp, C. (2007) The acquisition of inductive constraints. Unpublished doctoral dissertation, Massachusetts Institute of Technology. [aBML]

Kemp, C., Perfors, A. & Tenenbaum, J. B. (2007) Learning overhypotheses with hierarchical Bayesian models. *Developmental Science* 10(3):307–21. [rBML]

Kemp, C. & Tenenbaum, J. B. (2008) The discovery of structural form. *Proceedings of the National Academy of Sciences of the United States of America* 105 (31):10687–92. [rBML]

Keramati, M., Dezfouli, A. & Piray, P. (2011) Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS Computational Biology* 7: e1002055. [aBML]

Khaligh-Razavi, S. M. & Kriegeskorte, N. (2014) Deep supervised, but not unsupervised, models may explain IT cortical representation. *PLoS Computational Biology* 10(11):e1003915. [aBML, NK]

Kidd, C., Piantadosi, S. T. & Aslin, R. N. (2012) The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS One* 7(5):e36399. [P-YO]

Kiddon, C., Zettlemoyer, L. & Choi, Y. (2016). Globally coherent text generation with neural checklist models. In: *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, Austin, Texas, November 1–5, 2016,* pp. 329–39. Association for Computational Linguistics. [rBML]

Kilner, J. M., Friston, K. J. & Frith, C. D. (2007) Predictive coding: An account of the mirror neuron system. *Cognitive Processing* 8(3):159–66. [aBML]

Kingma, D. P., Rezende, D. J., Mohamed, S. & Welling, M. (2014) Semi-supervised learning with deep generative models. Presented at the 2014 Neural Information Processing Systems conference, Montreal, QC, Canada. In: *Advances in neural information processing systems 27 (NIPS 2014),* ed. Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence & K. Q. Weinberger [spotlight]. Neural Information Processing Systems Foundation. [aBML]

Kiraly, I., Csibra, G. & Gergely, G. (2013) Beyond rational imitation: Learning arbitrary means actions from communicative demonstrations. *Journal of Experimental Child Psychology* 116(2):471–86. [DCD]

Kline, M. A. (2015) How to learn about teaching: An evolutionary framework for the study of teaching behavior in humans and other animals. *Behavioral and Brain Sciences* 2015;38:e31. [JMC, MHT]

Koch, G., Zemel, R. S. & Salakhutdinov, R. (2015) Siamese neural networks for one-shot image recognition. Presented at the Deep Learning Workshop at the 2015 International Conference on Machine Learning, Lille, France. Available at: https://www.cs.cmu.edu/~rsalakhu/papers/oneshot1.pdf. [aBML]

Kodratoff, Y. & Michalski, R. S. (2014) *Machine: earning: An artificial intelligence approach, vol. 3*. Morgan Kaufmann. [aBML]

Kolodner, J. (1993) *Case-based reasoning*. Morgan Kaufmann. [NC]

Koza, J. R. (1992) *Genetic programming: On the programming of computers by means of natural selection, vol. 1*. MIT press. [aBML]

Kriegeskorte, N. (2015) Deep neural networks: A new framework for modeling biological vision and brain information processing. *Annual Review of Vision Science* 1:417–46. [aBML]

Kriegeskorte, N. & Diedrichsen, J. (2016) Inferring brain-computational mechanisms with models of activity measurements. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* 371(1705):489–95. [NK]

Kriegeskorte, N., Mur, M. & Bandettini, P. (2008) Representational similarity analysis – Connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience* 2:4. doi: 10.3389/neuro.06.004.2008. [NK]

Krizhevsky, A., Sutskever, I. & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. Presented at the 25th International Conference on Neural Information Processing Systems, Lake Tahoe, NV, December 3–6, 2012. In: *Advances in Neural Information Processing Systems 25 (NIPS 2012),* ed. F. Pereira, C. J. C. Burges, L. Bottou & K. Q. Weinberger, pp. 1097–105. Neural Information Processing Systems Foundation. [arBML, MB, NK, SSH]

Kulkarni, T. D., Kohli, P., Tenenbaum, J. B. & Mansinghka, V. (2015a) Picture: A probabilistic programming language for scene perception. In: *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, June 7–12, 2015*, pp. 4390–99. IEEE. [aBML]

Kulkarni, T. D., Narasimhan, K. R., Saeedi, A. & Tenenbaum, J. B. (2016) Hierarchical deep reinforcement learning: Integrating temporal abstraction and intrinsic motivation. *arXiv preprint 1604.06057*. Available at: https://arxiv.org/abs/1604.06057. [aBML, P-YO]

Kulkarni, T. D., Whitney, W., Kohli, P. & Tenenbaum, J. B. (2015b) Deep convolutional inverse graphics network. *arXiv preprint 1503.03167.* Available at: https://arxiv.org/abs/1503.03167. [aBML]

Lake, B. M. (2014) *Towards more human-like concept learning in machines: Compositionality, causality, and learning-to-learn*. Unpublished doctoral dissertation, Massachusetts Institute of Technology. [aBML]

Lake, B. M., Lawrence, N. D. & Tenenbaum, J. B. (2016) The emergence of organizing structure in conceptual representation. *arXiv preprint 1611.09384*. Available at: http://arxiv.org/abs/1611.09384. [MB, rBML]

Lake, B. M., Lee, C.-Y., Glass, J. R. & Tenenbaum, J. B. (2014) One-shot learning of generative speech concepts. In: *Proceedings of the 36th Annual Conference of the Cognitive Science Society, Quebec City, QC, Canada, July 23–26, 2014,* pp. 803–08. Cognitive Science Society. [aBML]

Lake, B. M., Salakhutdinov, R. & Tenenbaum, J. B. (2012) Concept learning as motor program induction: A large-scale empirical study. In: *Proceedings of the 34th Annual Conference of the Cognitive Science Society, Sapporo, Japan, August 1–4, 2012*, pp. 659–64. Cognitive Science Society. [aBML]

Lake, B. M., Salakhutdinov, R. & Tenenbaum, J. B. (2015a) Human-level concept learning through probabilistic program induction. *Science* 350(6266):1332–38. [arBML, MB, ED, NK]

Lake, B. M., Zaremba, W., Fergus, R. & Gureckis, T. M. (2015b) Deep neural networks predict category typicality ratings for images. In: *Proceedings of the 37th Annual Meeting of the Cognitive Science Society, Pasadena, CA, July 22–25, 2015*. Cognitive Science Society. ISBN: 978-0-9911967-2-2. [aBML]

Lakoff, G. & Johnson, M. (2003) *Metaphors we live by*, 2nd ed. University of Chicago Press. [SW]

Lambert, A (2011) *The gates of hell: Sir John Franklin's tragic quest for the Northwest Passage*. Yale University Press. [MHT]

Landau, B., Smith, L. B. & Jones, S. S. (1988) The importance of shape in early lexical learning. *Cognitive Development* 3(3):299–321. [aBML]

Lande, T. S., ed. (1998) Neuromorphic systems engineering: Neural networks in silicon. *Kluwer International Series in Engineering and Computer Science, vol. 447*. Kluwer Academic. ISBN 978-0-7923-8158-7. [KBC]

Langley, P., Bradshaw, G., Simon, H. A. & Zytkow, J. M. (1987) *Scientific discovery: Computational explorations of the creative processes*. MIT Press. [aBML]

Laptev, I., Marszalek, M., Schmid, C. & Rozenfeld, B. (2008) Learning realistic human actions from movies. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, June 23–28, 2008 (CVPR 2008)*, pp. 1–8. IEEE. [SW]

Larson, H. J., Cooper, L. Z., Eskola, J., Katz, S. L. & Ratzan, S. (2011) Addressing the vaccine confidence gap. *The Lancet* 378(9790):526–35. [EJL]

Lázaro-Gredilla, M., Liu, Y., Phoenix, D. S. & George, D. (2016) Hierarchical compositional feature learning. *arXiv preprint 1611.02252*. Available at: http://arxiv.org/abs/1611.02252. [DGe]

LeCun, Y., Bengio, Y. & Hinton, G. (2015) Deep learning. *Nature* 521:436–44. [aBML]

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W. & Jackel, L. D. (1989) Backpropagation applied to handwritten zip code recognition. *Neural Computation* 1:541–51. [arBML]

LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P. (1998) Gradient-based learning applied to document recognition. *Proceedings of the IEEE* 86(11):2278–323. [aBML]

Lee, T. S. (2015) The visual system's internal model of the world. *Proceedings of the IEEE* 103(8):1359–78. Available at: http://doi.org/10.1109/JPROC.2015.2434601. [DGe]

Legare, C. H. & Harris, P. L. (2016) The ontogeny of cultural learning. *Child Development* 87(3):633–42. [JMC]

Lenat, D. & Guha, R. V. (1990) *Building large. Knowledge based systems: Representation and inference in the Cyc project*. Addison-Wesley. [LRC]

Lenat, D., Miller, G. & Yokoi, T (1995) CYC, WordNet, and EDR: Critiques and responses. *Communications of the ACM* 38(11):45–48. [LRC]

Lerer, A., Gross, S. & Fergus, R. (2016) Learning physical intuition of block towers by example. Presented at the 33rd International Conference on Machine Learning. *Proceedings of Machine Learning Research* 48:430–08. [aBML]

Levy, R. P., Reali, F. & Griffiths, T. L. (2009) Modeling the effects of memory on human online sentence processing with particle filters. Presented at the 2008 Neural Information Processing Systems conference. Vancouver, BC, Canada, December 8–10, 2008. In: *Advances in neural information processing systems 21 (NIPS 2008)*, pp. 937–44. Neural Information Processing Systems. [aBML]

Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N. & Cook, J. (2012) Misinformation and its correction continued influence and successful debiasing. *Psychological Science in the Public Interest* 13(3):106–31. [EJL]

Lewis-Kraus, G. (2016) Going neural. *New York Times Sunday Magazine* 40–49+, December 18, 2016. [DEM]

Liang, C. and Forbus, K. (2015) Learning plausible inferences from semantic web knowledge by combining analogical generalization with structured logistic regression. In: *Proceedings of the 29th AAAI Conference on Artificial Intelligence, Austin, TX*. AAAI Press. [KDF]

Liao, Q., Leibo, J. Z. & Poggio, T. (2015) How important is weight symmetry in backpropagation? *arXiv preprint arXiv:1510.05067*. Available at: https://arxiv.org/abs/1510.05067. [aBML]

Liberman, A. M., Cooper, F. S., Shankweiler, D. P. & Studdert-Kennedy, M. (1967) Perception of the speech code. *Psychological Review* 74(6):431–61. [aBML]

Lillicrap, T. P., Cownden, D., Tweed, D. B. & Akerman, C. J. (2014) Random feedback weights support learning in deep neural networks. *arXiv preprint:1411.0247*. Available at: https://arxiv.org/abs/1411.0247. [aBML]

Lindeman, M. (2011) Biases in intuitive reasoning and belief in complementary and alternative medicine. *Psychology and Health* 26(3):371–82. [EJL]

Lisman, J. E. & Grace, A. A. (2005) The hippocampal-VTA loop: Controlling the entry of information into long-term memory. *Neuron* 46:703–13. [GB]

Liu, D., Wellman, H. M., Tardif, T., & Sabbagh, M. A. (2008). Theory of mind development in Chinese children: A meta-analysis of false-belief understanding across cultures and languages. *Developmental Psychology* 44(2):523–31. Available at: http://dx.doi.org/10.1037/0012-1649.44.2.523. [JMC, rBML]

Lloyd, J., Duvenaud, D., Grosse, R., Tenenbaum, J. & Ghahramani, Z. (2014) Automatic construction and natural-language description of nonparametric regression models. In: *Proceedings of the national conference on artificial intelligence* 2:1242–50. [aBML]

Logan, G. D. (1988) Toward an instance theory of automatization. *Psychological Review* 95(4):492–527. [NC]

Lombrozo, T. (2009) Explanation and categorization: How "why?" informs "what?". *Cognition* 110(2):248–53. [aBML]

Lombrozo, T. (2016) Explanatory preferences shape learning and inference. *Trends in Cognitive Sciences* 20(10):748–59. [rBML]

Lopes, M. & Oudeyer, P.-Y. (2012) The strategic student approach for life-long exploration and learning. In: *IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL)*, San Diego, CA, November 7–9, 2012, pp. 1–8. IEEE. [P-YO]

Lopes, M. & Santos-Victor, J. (2007). A developmental roadmap for learning by imitation in robots. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* 37(2):308–21. [LM]

Lopez-Paz, D., Bottou, L., Scholköpf, B. & Vapnik, V. (2016) Unifying distillation and privileged information. Presented at the International Conference on Learning Representations (ICLR), San Juan, Puerto Rico, May 2–4, 2016. *arXiv preprint 1511.03643v3*. Available at: https://arxiv.org/abs/1511.03643. [aBML]

Lopez-Paz, D., Muandet, K., Scholköpf, B. & Tolstikhin, I. (2015) Towards a learning theory of cause-effect inference. Presented at the 32nd International Conference on Machine Learning (ICML), Lille, France, July 7–9, 2015. *Proceedings of Machine Learning Research* 37:1452–61. [aBML]

Loughnan, S. & Haslam, N. (2007) Animals and androids implicit associations between social categories and nonhumans. *Psychological Science* 18(2):116–21. [rBML]

Lovett, A. & Forbus, K. (2017) Modeling visual problem solving as analogical reasoning. *Psychological Review* 124(1):60–90. [KDF]

Lungarella, M., Metta, G., Pfeifer, R. & Sandini, G. (2003) Developmental robotics: A survey. *Connection Science* 15:151–90. [BJM]

Luong, M.-T., Le, Q. V., Sutskever, I., Vinyals, O. & Kaiser, L. (2015) Multi-task sequence to sequence learning. *arXiv preprint 1511.06114*. Available at: https://arxiv.org/pdf/1511.06114.pdf. [aBML]

Lupyan, G. & Bergen, B. (2016) How language programs the mind. *Topics in Cognitive Science* 8(2):408–24. [aBML]

Lupyan, G. & Clark, A. (2015) Words and the world: Predictive coding and the language perception-cognition interface. *Current Directions in Psychological Science* 24(4):279–84. [aBML]

Macindoe, O. (2013) Sidekick agents for sequential planning problems. Unpublished doctoral dissertation, Massachusetts Institute of Technology. [aBML]

Mackenzie, D. (2012) A flapping of wings. *Science* 335(6075):1430–33. [DEM]

Magid, R. W., Sheskin, M. & Schulz, L. E. (2015) Imagination and the generation of new ideas. *Cognitive Development* 34:99–110. [aBML]

Mahoor, Z. MacLennan, B. & MacBride, A. (2016) Neurally plausible motor babbling in robot reaching. In: *The 6th Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics, September 19–22, 2016, Cergy-Pontoise/Paris,* pp. 9–14. IEEE. [BJM]

Malle, B. F. & Scheutz, M. (2014) Moral competence in social robots. In: *Proceedings of the 2014 IEEE International Symposium on Ethics in Science, Technology and Engineering*. IEEE. doi:10.1109/ETHICS.2014.6893446. [PMP]

Mannella, F. & Baldassarre, G. (2015) Selection of cortical dynamics for motor behaviour by the basal ganglia. *Biological Cybernetics* 109:575–95. [GB]

Mannella, F., Gurney, K. & Baldassarre, G. (2013) The nucleus accumbens as a nexus between values and goals in goal-directed behavior: A review and a new hypothesis. *Frontiers in Behavioral Neuroscience* 7(135):e1–29. [GB]

Mansinghka, V., Selsam, D. & Perov, Y. (2014) Venture: A higher-order probabilistic programming platform with programmable inference. *arXiv preprint 1404.0099*. Available at: https://arxiv.org/abs/1404.0099 [aBML]

Marblestone, A. H., Wayne, G. & Kording, K. P. (2016) Toward an integration of deep learning and neuroscience. *Frontiers in Computational Neuroscience* 10:94. [AHM, NK]

Marcus, G. (1998) Rethinking eliminative connectionism. *Cognitive Psychology* 282 (37):243–82. [aBML]

Marcus, G. (2001) *The algebraic mind: Integrating connectionism and cognitive science*. MIT Press. [aBML]

Marin, L., Issartel, J. & Chaminade, T. (2009). Interpersonal motor coordination: From human-human to human-robot interactions. *Interaction Studies* 10 (3):479–504. [LM]

Markman, A. B. & Makin, V. S. (1998) Referential communication and category acquisition. *Journal of Experimental Psychology: General* 127(4):331–54. [aBML]

Markman, A. B. & Ross, B. H. (2003) Category use and category learning. *Psychological Bulletin* 129(4):592–613. [aBML]

Markman, E. M. (1989) *Categorization and naming in children*. MIT Press. [aBML]

Marr, D. (1982/2010). *Vision*. MIT Press. [ESS]

Marr, D. (1982) *Vision: A computational investigation into the human representation and processing of visual information*. MIT Press. [SSH]

Marr, D. (1983) *Vision*. W. H. Freeman. [KDF]

Marr, D. C. (1982) *Vision*. W. H. Freeman. [aBML]

Marr, D. C. & Nishihara, H. K. (1978) Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London Series B: Biological Sciences* 200(1140):269–94. [aBML]

Mascalzoni, E., Regolin, L. & Vallortigara, G. (2010). Innate sensitivity for self-propelled causal agency in newly hatched chicks. *Proceedings of the National Academy of Sciences of the United States of America* 107(9):4483–85. [ESS]

Maslow, A. (1954) *Motivation and personality*. Harper & Brothers. [CDG]

Matute, H., Blanco, F., Yarritu, I., Díaz-Lago, M., Vadillo, M. A. & Barberia, I. (2015) Illusions of causality: How they bias our everyday thinking and how they could be reduced. *Frontiers in Psychology* 6:888. doi: 10.3389/fpsyg.2015.00888. [EJL]

Mayer, J. D. & Salovey, P. (1993) The intelligence of emotional intelligence. *Intelligence* 17:442–43. [RJS]

Mazur, J. E. & Hastie, R. (1978) Learning as accumulation: A reexamination of the learning curve. *Psychological Bulletin* 85:1256–74. [LRC]

McCarthy, J. (1959) Programs with common sense at the Wayback machine (archived October 4, 2013). In: *Proceedings of the Teddington Conference on the Mechanization of Thought Processes*, pp. 756–91. AAAI Press. [LRC]

McCarthy, J. & Hayes, P. J. (1969) Some philosophical problems from the standpoint of artificial intelligence. In: Machine Intelligence 4, ed. B. Meltzer & D. Michie, pp. 463–502. Edinburgh University Press. [LRC]

McClelland, J. L. (1988) *Parallel distributed processing: Implications for cognition and development* [technical report]. Defense Technical Information Center document. Available at: http://www.dtic.mil/get-tr-doc/pdf?AD=ADA219063. [aBML]

McClelland, J. L., Botvinick, M. M., Noelle, D. C., Plaut, D. C., Rogers, T. T., Seidenberg, M. S. & Smith, L. B. (2010) Letting structure emerge: Connectionist and dynamical systems approaches to cognition. *Trends in Cognitive Sciences* 14(8):348–56. [arBML]

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. *Psychological Review* 102(3):419–57. [arBML]

McClelland, J. L. & Rumelhart, D. E. (1986) *Parallel distributed processing: Explorations in the microstructure of cognition, Vol. 2.* MIT Press. [aBML]

McFate, C. & Forbus, K. (2016) An analysis of frame semantics of continuous processes. *Proceedings of the 38th Annual Conference of the Cognitive Science Society, Philadelphia, PA,* ed. A. Papafragou, D. Grodner, D. Mirman & J. C. Trueswell, pp. 836–41. Cognitive Science Society. [KDF]

McFate, C. J., Forbus, K. & Hinrichs, T. (2014) Using narrative function to extract qualitative information from natural language texts. *Proceedings of the 28th AAAI Conference on Artificial Intelligence, Québec City, Canada, July 27–31, 2014*, pp. 373–379. AAAI Press. [KDF]

McShea, D. W. (2013) Machine wanting. *Studies on the History and Philosophy of Biological and Biomedical Sciences* 44(4 pt B):679–87. [KBC]

Medin, D. L. & Ortony, A. (1989). Psychological essentialism. In: *Similarity and analogical reasoning*, ed. S. Vosniadou & A. Ortony, pp. 179–95. Cambridge University Press. [rBML]

Medin, D. L. & Schaffer, M. M. (1978) Context theory of classification learning. *Psychological Review* 85(3):207–38. [NC]

Mejía-Arauz, R., Rogoff, B. & Paradise, R. (2005) Cultural variation in children's observation during a demonstration. *International Journal of Behavioral Development* 29(4):282–91. [JMC]

Meltzoff, A. N. (2007). 'Like me': A foundation for social cognition. *Developmental Science* 10(1):126–34. [LM]

Meltzoff, A. N., Kuhl, P. M., Movellan, J. & Sejnowski, T. J. (2009) Foundations for a new science of learning. *Science* 325(5938):284–88. [KBC]

Meltzoff, A. N. & Moore, M. K. (1995) Infants' understanding of people and things: From body imitation to folk psychology. In: *The body and the self*, ed. J. L. Bermúdez, A. Marcel & N. Eilan, pp. 43–70. MIT Press. [JMC]

Meltzoff, A. N. & Moore, M. K. (1997) Explaining facial imitation: a theoretical model. *Early Development and Parenting* 6:179–92. [BJM]

Mesoudi, A., Chang, L., Murray, K. & Lu, H. J. (2015) Higher frequency of social learning in China than in the West shows cultural variation in the dynamics of

cultural evolution. *Proceeding of the Royal Society of London Series B: Biological Sciences* 282(1798):20142209. [JMC]
Metcalfe, J., Cottrell, G. W. & Mencl, W. E. (1992) Cognitive binding: A computational-modeling analysis of a distinction between implicit and explicit memory. *Journal of Cognitive Neuroscience* 4(3):289–98. [LRC]
Mikolov, T., Joulin, A. & Baroni, M. (2016) A roadmap towards machine intelligence. *arXiv preprint 1511.08130*. Available at: http://arxiv.org/abs/1511.08130. [arBML]
Mikolov, T., Sutskever, I. & Chen, K. (2013) Distributed representations of words and phrases and their compositionality. Presented at the 2013 Neural Information Processing Systems conference, Lake Tahoe, NV, December 5–10, 2013. In: *Advances in Neural Information Processing Systems 26 (NIPS)*, ed C. J. C. Burges, L. Bottou, M. Welling, Z. Ghagramani & K. Q. Weinberger [poster].Neural Information Processing Systems Foundation. [aBML]
Miller, E. G., Matsakis, N. E. & Viola, P. A. (2000) Learning from one example through shared densities on transformations. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head Island, SC, June 15, 2000*. IEEE. [aBML]
Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D. & Miller, K. J. (1990) Introduction to WordNet: An on-line lexical database. *International Journal of Lexicography* 3(4):235–44. [LRC]
Miller, G. A. & Johnson-Laird, P. N. (1976) *Language and perception*. Belknap Press. [aBML]
Milner, D. & Goodale, M. (2006) *The visual brain in action*. Oxford University Press. [GB]
Minsky, M. (1986) *The society of mind*. Simon and Schuster. [MBu]
Minsky, M. (2003) *Semantic information processing.* MIT Press. [RJS]
Minsky, M. & Papert, S. A. (1987) *Perceptrons: An introduction to computational geometry*, expanded edn. MIT Press. [RJS]
Minsky, M. L. (1974) A framework for representing knowledge. MIT-AI Laboratory Memo 306. [aBML]
Minsky, M. L. & Papert, S. A. (1969) *Perceptrons: An introduction to computational geometry*. MIT Press. [aBML]
Mirolli, M., Mannella, F. & Baldassarre, G. (2010) The roles of the amygdala in the affective regulation of body, brain and behaviour. *Connection Science* 22 (3):215–45. [GB]
Mišić, B., Sporns, O. & McIntosh, A. R. (2014) Communication efficiency and congestion of signal traffic in large-scale brain networks. *PLoS Computational Biology* 10(1):e1003427. [DG]
Mitchell, T. M., Keller, R. R. & Kedar-Cabelli, S. T. (1986) Explanation-based generalization: A unifying view. *Machine Learning* 1:47–80. [aBML]
Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S. & Floridi, L. (2016) The ethics of algorithms: Mapping the debate. *Big Data & Society* 3(2):1–21. doi:10.1177/2053951716679679. [PMP]
Mix, K. S. (1999) Similarity and numerical equivalence: Appearances count. *Cognitive Development* 14:269–97. [KDF]
Mnih, A. & Gregor, K. (2014) Neural variational inference and learning in belief networks. Presented at the 31st International Conference on Machine Learning, Beijing, China, June 22–24, 2014. *Proceedings of Machine Learning Research* 32:1791–99. [aBML]
Mnih, V., Heess, N., Graves, A. & Kavukcuoglu, K. (2014). Recurrent models of visual attention. Presented at the 28th Annual Conference on Neural Information Processing Systems, Montreal, Canada. In: *Advances in Neural Information Processing Systems* 27*(NIPS 2014)*, ed. Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence & K. Q. Weinberger. Neural Information Processing Systems Foundation. [arBML]
Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D. & Riedmiller, M. (2013) Playing Atari with deep reinforcement learning. *arXiv preprint 1312.5602*. Available at: https://arxiv.org/abs/1312.5602. [SSH]
Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglous, I., King, H., Kumaran, D., Wierstra, D. & Hassabis, D. (2015) Human-level control through deep reinforcement learning. *Nature* 518(7540):529–33. [arBML, MB, DGe]
Moeslund, T. B., Hilton, A. & Krüger, V. (2006) A survey of advances in vision-based human motion capture and analysis. *Computer Vision and Image Understanding* 104(2):90–126. [rBML]
Mogenson, G. J., Jones, D. L. & Yim, C. Y. (1980) From motivation to action: Functional interface between the limbic system and the motor system. *Progress in Neurobiology* 14(2–3):69–97. [GB]
Mohamed, S. & Rezende, D. J. (2015) Variational information maximisation for intrinsically motivated reinforcement learning. Presented at the 2015 Neural Information Processing Systems conference, Montreal, QC, Canada, December 7–12, 2015. *Advances in Neural Information Processing Systems 28 (NIPS 2015)*, ed. C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama & R. Garnett, pp. 2125–33. Neural Information Processing Systems Foundation. [aBML]
Moreno-Bote, R., Knill, D. C. & Pouget, A. (2011) Bayesian sampling in visual perception. *Proceedings of the National Academy of Sciences of the United States of America* 108:12491–96. [aBML]
Moser, E., Kropff, E. & Moser, M. B. (2008). Place cells, grid cells, and the brain's spatial representation system. *Annual Review of Neuroscience* 31:69–89. [ESS]
Moulin-Frier, C., Nguyen, M. & Oudeyer, P.-Y. (2014) Self-organization of early vocal development in infants and machines: The role of intrinsic motivation. *Frontiers in Psychology* 4:1006. Available at: http://dx.doi.org/10.3389/fpsyg.2013.01006. [P-YO]
Murphy, G. L. (1988) Comprehending complex concepts. *Cognitive Science* 12 (4):529–62. [aBML]
Murphy, G. L. & Medin, D. L. (1985) The role of theories in conceptual coherence. *Psychological Review* 92(3):289–316. [arBML]
Murphy, G. L. & Ross, B. H. (1994) Predictions from uncertain categorizations. *Cognitive Psychology* 27:148–93. [aBML]
Nagai, Y., Kawai, Y. & Asada, M. (2011). Emergence of mirror neuron system: Immature vision leads to self-other correspondence. *Proceedings of the 1st Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics, Vol. 2*, pp. 1–6. IEEE. [LM]
Nakayama, K., Shimojo, S. & Silverman, G. H. (1989) Stereoscopic depth: Its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception* 18:55–68. [rBML]
Neisser, U. (1966) *Cognitive psychology*. Appleton-Century-Crofts. [aBML]
Newell, A. (1990) *Unified theories of cognition.* Harvard University Press. [RPK]
Newell, A., Shaw, J. C. & Simon, H. A. (1957) Problem solving in humans and computers. *Carnegie Technical* 21(4):34–38. [RJS]
Newell, A. & Simon, H. (1956) The logic theory machine. A complex information processing system. *IRE Transactions on Information Theory* 2(3):61–79. [LRC]
Newell, A. & Simon, H. A. (1961) *GPS, A program that simulates human thought*. Defense Technical Information Center. [aBML]
Newell, A. & Simon, H. A. (1972) *Human problem solving*. Prentice-Hall. [aBML]
Nguyen, M. & Oudeyer, P.-Y. (2013) Active choice of teachers, learning strategies and goals for a socially guided intrinsic motivation learner, *Paladyn Journal of Behavioural Robotics* 3(3):136–46. [P-YO]
Nguyen-Tuong, D. & Peters, J. (2011) Model learning for robot control: A survey. *Cognitive Processing* 12(4):319–40. [P-YO]
Nisbett, R. E. & Ross, L. (1980) *Human inference: Strategies and shortcomings of social judgment.* Prentice-Hall. ISBN 0-13-445073-6. [KBC, NC]
Niv, Y. (2009) Reinforcement learning in the brain. *Journal of Mathematical Psychology* 53:139–54. [aBML]
Norman, D. A. & Shallice, T. (1986) Attention to action: willed and automatic control of behaviour. In: *Advances in research: Vol. IV. Consciousness and self regulation*, ed. R. Davidson, G. Schwartz & D. Shapiro. Plenum. [RPK]
Oaksford, M. & Chater, N. (1991) Against logicist cognitive science. *Mind and Language* 6(1):1–38. [NC]
O'Donnell, T. J. (2015) *Productivity and reuse in language: A theory of linguistic computation and storage*. MIT Press. [aBML]
O'Keefel (2014). *Nobel lecture: Spatial cells in the hippocampal formation*. Available at: http://www.nobelprize.org/nobel_prizes/medicine/laureates/2014/okeefe-lecture.html. [ESS]
O'Keefe, J. & Nadel, L. (1978). *The hippocampus as a cognitive map.* Oxford University Press. [ESS]
Olshausen, B. A., Anderson, C. H. & Van Essen, D. C. (1993) A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *The Journal of Neuroscience* 13(11):4700–19. [DG]
Ong, D. C., Zaki, J. & Goodman, N. D. (2015) Affective cognition: Exploring lay theories of emotion. *Cognition* 143:141–62. [rBML]
O'Regan, J.K. (2011). *Why red doesn't sound like a bell: Understanding the feel of consciousness.* Oxford University Press. [LM]
Osherson, D. N. & Smith, E. E. (1981) On the adequacy of prototype theory as a theory of concepts. *Cognition* 9(1):35–58. [aBML]
Otto, A. R., Skatova, A., Madlon-Kay, S. & Daw, N. D. (2015) Cognitive control predicts use of model-based reinforcement learning. *Journal of Cognitive Neuroscience* 27:319–33. [EJL]
Oudeyer, P.-Y. (2016) What do we learn about development from baby robots? *WIREs Cognitive Science* 8(1–2):e1395. Available at: http://www.pyoudeyer.com/oudeyerWiley16.pdf. doi: 10.1002/wcs.1395. [P-YO]
Oudeyer, P.-Y., Baranes, A. & Kaplan, F. (2013) Intrinsically motivated learning of real-world sensorimotor skills with developmental constraints. In: *Intrinsically motivated learning in natural and artificial systems*, ed. G. Baldassarre & M. Mirolli, pp. 303–65. Springer. [P-YO]
Oudeyer, P.-Y., Kaplan, F. & Hafner, V. (2007) Intrinsic motivation systems for autonomous mental development. *IEEE Transactions on Evolutionary Computation* 11(2):265–86. [P-YO]
Oudeyer, P.-Y. & Smith, L. (2016) How evolution may work through curiosity-driven developmental process. *Topics in Cognitive Science* 8(2):492–502. [P-YO]
Palmer, S. (1999) *Vision science: Photons to phenomenology*. MIT Press. [KDF]
Parisotto, E., Ba, J. L. & Salakhutdinov, R. (2016) Actor-mimic: Deep multitask and transfer reinforcement learning. Presented at the International Conference on

Learning Representations (ICLR), San Juan, Puerto Rico. May 2–5, 2016. *arXiv preprint 1511.06342v4*. Available at: https://www.google.com/search?q=arXiv%3A+preprint+1511.06342v4&ie=utf-8&oe=utf-8. [aBML]

Parker, S. T. & McKinney, M. L. (1999) *Origins of intelligence: The evolution of cognitive development in monkeys, apes and humans*. Johns Hopkins University Press. ISBN 0-8018-6012-1. [KBC]

Pecevski, D., Buesing, L. & Maass, W. (2011) Probabilistic inference in general graphical models through sampling in stochastic networks of spiking neurons. *PLoS Computational Biology* 7:e1002294. [aBML]

Penhune, V. B. & Steele, C. J. (2012) Parallel contributions of cerebellar, striatal and M1 mechanisms to motor sequence learning. *Behavioural Brain Research* 226 (2):579–91. [GB]

Peterson, J. C., Abbott, J. T. & Griffiths, T. L. (2016) Adapting deep network features to capture psychological representations. In: *Proceedings of the 38th Annual Conference of the Cognitive Science Society, Philadelphia, Pennsylvania, August 10–13, 2016,* ed. A. Papafragou, Daniel J. Grodner, D. Mirman & J. Trueswell, pp. 2363–68. Cognitive Science Society. [aBML]

Pfeifer, R. & Gómez, G. (2009) Morphological computation–connecting brain, body, and environment. In: *Creating brain-like intelligence*, ed. B. Sendhoff, E Körner, H. Ritter & K. Doya, pp. 66–83. Springer. [GB]

Pfeifer, R., Lungarella, M. & Iida, F. (2007) Self-organization, embodiment, and biologically inspired robotics. *Science* 318(5853):1088–93. [P-YO]

Piantadosi, S. T. (2011) Learning and the language of thought. Unpublished doctoral dissertation, Massachusetts Institute of Technology. [aBML]

Pinker, S. (2007) *The stuff of thought: Language as a window into human nature*. Penguin. [aBML]

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

Poggio, T. (1984) Routing thoughts. Massachusetts Institute of Technology Artificial Intelligence Laboratory Working Paper 258. [DG]

Power, J. M., Thompson, L. T., Moyer, J. R. & Disterhoft, J. F. (1997) Enhanced synaptic transmission in ca1 hippocampus after eyeblink conditioning. *Journal of Neurophysiology* 78:1184–87. [aBML]

Prasada, S. & Pinker, S. (1993) Generalizations of regular and irregular morphology. *Language and Cognitive Processes* 8(1):1–56. [LRC]

Pratt, G. (2016, December 6). Presentation to Professor Deb Roy's class on machine learning and society at the MIT Media Lab. Class presentation that was videotaped but has not been made public. [DCD]

Premack, D. & Premack, A. J. (1997) Infants attribute value to the goal-directed actions of self-propelled objects. *Cognitive Neuroscience* 9(6):848–56. doi: 10.1162/jocn.1997.9.6.848. [aBML]

Putnam, H. (1967) Psychophysical predicates. In: *Art, mind, and religion*, ed. W. Capitan & D. Merrill. University of Pittsburgh Press. (Reprinted in 1975 as *The nature of mental states*, pp. 429–40. Putnam.) [LRC]

Ranzato, M., Szlam, A., Bruna, J., Mathieu, M., Collobert, R. & Chopra, S. (2016) Video (language) modeling: A baseline for generative models of natural videos. *arXiv preprint 1412.6604.* Available at: https://www.google.com/search?q=arXiv+preprint+1412.6604&ie=utf-8&oe=utf-8. [MB]

Raposo, D., Santoro, A., Barrett, D. G. T., Pascanu, R., Lillicrap, T. & Battaglia, P. (2017) Discovering objects and their relations from entangled scene representations. Presented at the Workshop Track at the International Conference on Learning Representations, Toulon, France, April 24–26, 2017. *arXiv preprint 1702.05068*. Available at: https://openreview.net/pdf?id=Bk2TqVcxe. [MB, rBML]

Ravi, S. & Larochelle, H. (2017) Optimization as a model for few-shot learning. Presented at the International Conference on Learning Representations, Toulon, France, April 24–26, 2017. Available at: https://openreview.net/pdf?id=rJY0-Kcll. [MB]

Read, S. J., Monroe, B. M., Brownstein, A. L., Yang, Y., Chopra, G. & Miller, L. C. (2010) A neural network model of the structure and dynamics of human personality. *Psychological Reviews* 117(1):61–92. [KBC]

Real, E., Moore, S., Selle, A., Saxena, S., Suematsu, Y. L., Le, Q. & Kurakin, A. (2017) Large-scale evolution of image classifiers. *arXiv preprint 1703.01041*. Available at: https://arxiv.org/abs/1703.01041. [rBML]

Redgrave, P. & Gurney, K. (2006) The short-latency dopamine signal: a role in discovering novel actions? *Nature Reviews Neuroscience* 7:967–75. [GB]

Reed, S. & de Freitas, N. (2016) Neural programmer-interpreters. Presented at the 4th International Conference on Learning Representations (ICLR), San Juan, Puerto Rico, May 2–5, 2016. *arXiv preprint 1511.06279*. Available at: https://arxiv.org/abs/1511.06279. [arBML, MB]

Regolin, L., Vallortigara, G. & Zanforlin, M. (1995). Object and spatial representations in detour problems by chicks. *Animal Behaviour* 49:195–99. [ESS]

Rehder, B. (2003) A causal-model theory of conceptual representation and categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29(6):1141–59. [aBML]

Rehder, B. & Hastie, R. (2001) Causal knowledge and categories: The effects of causal beliefs on categorization, induction, and similarity. *Journal of Experimental Psychology: General* 130(3):323–60. [aBML]

Rehling, J. A. (2001) Letter spirit (part two): Modeling creativity in a visual domain. Unpublished doctoral dissertation, Indiana University. [aBML]

Rezende, D. J., Mohamed, S., Danihelka, I., Gregor, K. & Wierstra, D. (2016) One-shot generalization in deep generative models. Presented at the International Conference on Machine Learning, New York, NY, June 20–22, 2016. *Proceedings of Machine Learning Research* 48:1521–29. [arBML, MB]

Rezende, D. J., Mohamed, S. & Wierstra, D. (2014) Stochastic backpropagation and approximate inference in deep generative models. Presented at the International Conference on Machine Learning (ICML), Beijing, China, June 22–24, 2014. *Proceedings of Machine Learning Research* 32:1278–86. [aBML]

Richland, L. E. & Simms, N. (2015) Analogy, higher order thinking, and education. *Wiley Interdisciplinary Reviews: Cognitive Science* 6(2):177–92. [KDF]

Rips, L. J. (1975) Inductive judgments about natural categories. *Journal of Verbal Learning and Verbal Behavior* 14(6):665–81. [aBML]

Rips, L. J. & Hespos, S. J. (2015) Divisions of the physical world: Concepts of objects and substances. *Psychological Bulletin* 141:786–811. [aBML]

Rock, I. (1983) *The logic of perception*. MIT Press. [NC]

Rogers, T. T. & McClelland, J. L. (2004) *Semantic cognition*. MIT Press. [aBML]

Rogoff, B. (2003) *The cultural nature of human development*. Oxford University Press. [JMC]

Rohlfing, K. J. & Nomikou, I. (2014) Intermodal synchrony as a form of maternal responsiveness: Association with language development. *Language, Interaction and Acquisition* 5(1):117–36. [SW]

Romanes, G. J. (1884) *Animal intelligence*. Appleton. [KBC]

Rosenblatt, F. (1958) The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review* 65:386–408. [aBML]

Rougier, N. P., Noelle, D. C., Braver, T. S., Cohen, J. D. & O'Reilly, R. C. (2005) Prefrontal cortex and flexible cognitive control: Rules without symbols. *Proceedings of the National Academy of Sciences of the United States of America* 102(20):7338–43. [aBML]

Rozenblit, L. & Keil, F. (2002) The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science* 26(5):521–62. [EJL, NC]

Ruciński, M. (2014) Modelling learning to count in humanoid robots. Ph.D. thesis, University of Plymouth, UK. [SW]

Rumelhart, D. E., Hinton, G. & Williams, R. (1986a) Learning representations by back-propagating errors. *Nature* 323(9):533–36. [aBML]

Rumelhart, D. E. & McClelland, J. L. (1986) On learning the past tenses of English verbs. In: *Parallel distributed processing: Explorations in the microstructure of cognition, Vol. 1*, ed. Rumelhart, D. F., McClelland, J. L. & PDP Research Group, pp. 216–71. MIT Press. [aBML]

Rumelhart, D. E., McClelland, J. L. & PDP Research Group. (1986b) *Parallel distributed processing: Explorations in the microstructure of cognition, Vol. 1.* MIT Press. [aBML]

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C. & Fei-Fei, L. (2015) ImageNet large scale visual recognition. *International Journal of Computer Vision* 115 (3):211–52. [aBML]

Russell, S. & Norvig, P. (2003) *Artificial intelligence: A modern approach*. Prentice–Hall. [aBML]

Rusu, A. A., Rabinowitz, N. C., Desjardins, G., Soyer, H., Kirkpatrick, J., Kavukcuoglu, K., Pascanu, R. & Hadsell, R. (2016) Progressive neural networks. *arXiv preprint 1606.04671*. Available at: http://arxiv.org/abs/1606.04671. [aBML]

Ryan, R. M. & Deci, E. L. (2007) Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemporary Educational Psychology* 25:54–67. [aBML]

Salakhutdinov, R., Tenenbaum, J. & Torralba, A. (2012) One-shot learning with a hierarchical nonparametric Bayesian model. *JMLR Workshop on Unsupervised and Transfer Learning* 27:195–207. [aBML]

Salakhutdinov, R., Tenenbaum, J. B. & Torralba, A. (2013) Learning with hierarchical-deep models. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35(8):1958–71. [aBML]

Salakhutdinov, R., Torralba, A. & Tenenbaum, J. (2011) Learning to share visual appearance for multiclass object detection. In: *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Colorado Springs, CO, June 20–25, 2011*, pp. 1481–88. IEEE. [aBML]

Sanborn, A. N., Mansingkha, V. K. & Griffiths, T. L. (2013) Reconciling intuitive physics and Newtonian mechanics for colliding objects. *Psychological Review* 120(2):411–37. [aBML, ED]

Santoro, A., Bartunov, S., Botvinick, M., Wierstra, D. & Lillicrap, T. (2016). Meta-learning with memory-augmented neural networks. Presented at the 33rd International Conference on Machine Learning, New York, NY, June 19–24, 2016. *Proceedings of Machine Learning Research* 48:1842–50. [MB, rBML]

Santoro, A., Bartunov, S., Botvinick, M., Wierstra, D., & Lillicrap, T. (2016). One-shot learning with memory-augmented neural networks. *arXiv preprint 1605.06065*. Available at: https://arxiv.org/abs/1605.06065. [SSH]

Santucci, V. G., Baldassarre, G. & Mirolli, M. (2016), GRAIL: A goal-discovering robotic architecture for intrinsically-motivated learning, *IEEE Transactions on Cognitive and Developmental Systems* 8(3):214–31. [GB]

Saxe, A. M., McClelland, J. L. & Ganguli, S. (2013) Dynamics of learning in deep linear neural networks. Presented at the NIPS 2013 Deep Learning Workshop, Lake Tahoe, NV, December 9, 2013. [LRC]

Saxe, A. M., McClelland, J. L. & Ganguli, S. (2014) Exact solutions to the nonlinear dynamics of learning in deep linear neural networks. Presented at the International Conference on Learning Representations, Banff, Canada, April 14–16, 2014. *arXiv preprint 1312.6120*. Available at:https://arxiv.org/abs/1312.6120. [LRC]

Scellier, B. & Bengio, Y. (2016) Towards a biologically plausible backprop. *arXiv preprint 1602.05179*. Available at: https://arxiv.org/abs/1602.05179v2. [aBML]

Schank, R. C. (1972) Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology* 3:552–631. [aBML]

Schaul, T., Quan, J., Antonoglou, I. & Silver, D. (2016) Prioritized experience replay. Presented at International Conference on Learning Representations (ICLR), San Diego, CA, May 7–9, 2015. *arXiv preprint 1511.05952*. Available at: https://arxiv.org/abs/1511.05952. [aBML, MB]

Schlegel, A., Alexander, P. & Peter, U. T. (2015) Information processing in the mental workspace is fundamentally distributed. *Journal of Cognitive Neuroscience* 28(2):295–307. [DG]

Schlottmann, A., Cole, K., Watts, R. & White, M. (2013) Domain-specific perceptual causality in children depends on the spatio-temporal configuration, not motion onset. *Frontiers in Psychology* 4:365. [aBML]

Schlottmann, A., Ray, E. D., Mitchell, A. & Demetriou, N. (2006) Perceived physical and social causality in animated motions: Spontaneous reports and ratings. *Acta Psychologica* 123:112–43. [aBML]

Schmidhuber, J. (1991) Curious model-building control systems. *Proceedings of the IEEE International Joint Conference on Neural Networks* 2:1458–63. [P-YO]

Schmidhuber, J. (2015) Deep learning in neural networks: An overview. *Neural Networks* 61:85–117. [aBML]

Schmidt, R.C. & Richardson, M.J. (2008). Dynamics of interpersonal coordination. In: *Coordination: Neural, behavioural and social dynamics,* ed. A. Fuchs & V. Jirsa, pp. 281–307. Springer-Verlag. [LM]

Scholl, B. J. & Gao, T. (2013) Perceiving animacy and intentionality: Visual processing or higher-level judgment? In: *Social perception: detection and interpretation of animacy, agency, and intention*, ed. M. D. Rutherford & V. A. Kuhlmeier. MIT Press Scholarship Online. [aBML]

Schuller, I. K., Stevens, R. & Committee Chairs (2015) *Neuromorphic computing: From materials to architectures. Report of a roundtable convened to consider neuromorphic computing basic research needs*. Office of Science, U.S. Department of Energy. [KBC]

Schultz, W., Dayan, P. & Montague, P. R. (1997) A neural substrate of prediction and reward. *Science* 275:1593–99. [aBML]

Schulz, L. (2012a) Finding new facts; thinking new thoughts. Rational constructivism in cognitive development. *Advances in Child Development and Behavior* 43:269–94. [rBML]

Schulz, L. (2012b) The origins of inquiry: Inductive inference and exploration in early childhood. *Trends in Cognitive Sciences* 16(7):382–89. [arBML]

Schulz, L. E., Gopnik, A. & Glymour, C. (2007) Preschool children learn about causal structure from conditional interventions. *Developmental Science* 10:322–32. [aBML]

Scott, R. (Director). (2007) Blade Runner: The Final Cut: Warner Brothers (original release, 1982). [DEM]

Scott, S. H. (2004) Optimal feedback control and the neural basis of volitional motor control. *Nature Reviews Neuroscience* 5(7):532–46. [GB]

Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R. & LeCun, Y. (2014) OverFeat: Integrated recognition, localization and detection using convolutional networks. Presented at the International Conference on Learning Representations (ICLR), Banff, Canada, April 14–16, 2014. *arXiv preprint 1312.6229v4*. Available at: https://arxiv.org/abs/1312.6229. [aBML]

Shadmehr, R. & Krakauer, J. W. (2008) A computational neuroanatomy for motor control. *Experimental Brain Research* 185(3):359–81. [GB]

Shafto, P., Goodman, N. D. & Frank, M. C. (2012) Learning from others: The consequences of psychological reasoning for human learning. *Perspectives on Psychological Science* 7(4):341–51. [MHT]

Shafto, P., Goodman, N. D. & Griffiths, T. L. (2014) A rational account of pedagogical reasoning: Teaching by, and learning from, examples. *Cognitive Psychology* 71:55–89. [aBML]

Shahaeian, A., Peterson, C. C., Slaughter, V. & Wellman, H. M. (2011) Culture and the sequence of steps in theory of mind development. *Developmental Psychology* 47(5):1239–47. [JMC]

Shallice, T. & Cooper, R. P. (2011) *The organisation of mind*. Oxford University Press. [RPK]

Shultz, T. R. (2003) *Computational developmental psychology*. MIT Press. [aBML]

Siegler, R. (1976) Three aspects of cognitive development. *Cognitive Psychology* 8 (4):481–520. [ED]

Siegler, R. S. & Chen, Z. (1998) Developmental differences in rule learning: A microgenetic analysis. *Cognitive Psychology* 36(3):273–310. [aBML]

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Driessche, G. V. D., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K, Graepel, T. & Hassabis, D. (2016) Mastering the game of go with deep neural networks and tree search. *Nature* 529(7585):484–89. [arBML, MB]

Silver, D., van Hasselt, H., Hessel, M., Schaul, T., Guez, A., Harley, T., Dulac-Arnold, G. Reichert, D., Rabinowitz, N., Barreto, A. & Degris, T. (2017) The predictron: End-to-end learning and planning. In: *Proceedings of the 34rd International Conference on Machine Learning, Sydney, Australia,* ed. M. F. Balcan & K. Q. Weinberger. [MB]

Silverman, R. D. & Hendrix, K. S. (2015) Point: Should childhood vaccination against measles be a mandatory requirement for attending school? Yes. *CHEST Journal* 148(4):852–54. [EJL]

Simon, H. A. (1967) Motivational and emotional controls of cognition. *Psychological Review* 74:29–39. [CDG]

Sizemore, A., Giusti, C., Betzel, R. F. & Bassett, D. S. (2016) Closures and cavities in the human connectome. *arXiv preprint 1608.03520*. Available at: https://arxiv.org/abs/1608.03520. [DG]

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L. & Samuelson, L. (2002) Object name learning provides on-the-job training for attention. *Psychological Science* 13(1):13–19. [aBML]

Socher, R., Perelygin, A., Wu, J. Y., Chuang, J., Manning, C. D., Ng, A. Y. & Potts, C. (2013) Recursive deep models for semantic compositionality over a sentiment treebank. In: *Proceedings of the Conference on EmpiricalMethods in Natural Language Processing (EMNLP), Seattle, WA*, vol. 1631, p. 1642. Association for Computational Linguistics. [rBML]

Solomon, K., Medin, D. & Lynch, E. (1999) Concepts do more than categorize. *Trends in Cognitive Sciences* 3(3):99–105. [aBML]

Spelke, E. S. (1990) Principles of object perception. *Cognitive Science* 14(1):29–56. [aBML, JMC]

Spelke, E. S. (2003) What makes us smart? Core knowledge and natural language. Spelke ES. What makes us smart? Core knowledge and natural language. In: *Language in mind: Advances in the Investigation of language and thought,* ed. D. Gentner & S. Goldin-Meadow, pp. 277–311. MIT Press. [arBML]

Spelke, E. S., Gutheil, G. & Van de Walle, G. (1995) The development of object perception. In: *An invitation to cognitive science: vol. 2. Visual cognition,* 2nd ed. pp. 297–330. Bradford. [aBML]

Spelke, E. S. & Kinzler, K. D. (2007) Core knowledge. *Developmental Science* 10 (1):89–96. [arBML]

Spelke, E. S. & Lee, S. A. (2012). Core systems of geometry in animal minds. *Philosophical Transactions of the Royal Society, B: Biological Sciences* 367 (1603):2784–93. [ESS]

Squire, L. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys and humans. *Psychological Review* 99(2):195–231. [ESS]

Srivastava, N. & Salakhutdinov, R. (2013) Discriminative transfer learning with tree-based priors. Presented at the 2013 Neural Information Processing Systems conference, Lake Tahoe, NV, December 5–10, 2013. In: *Advances in Neural Information Processing Systems 26 (NIPS 2013)*, ed. C J. C. Burges, L. Bottou, M. Welling, Z. Ghagramani & K. Q. Weinberger [poster]. Neural Information Processing Systems Foundation. [aBML]

Stadie, B. C., Levine, S. & Abbeel, P. (2016) Incentivizing exploration in reinforcement learning with deep predictive models. *arXiv preprint 1507.00814*. Available at: http://arxiv.org/abs/1507.00814. [aBML]

Stahl, A. E. & Feigenson, L. (2015) Observing the unexpected enhances infants' learning and exploration. *Science* 348(6230):91–94. [aBML]

Stanley, K. O. & Miikkulainen, R. (2002) Evolving neural networks through augmenting topologies. *Evolutionary Computation* 10(2):99–127. [rBML]

Stanovich, K. E. (2009) *What intelligence tests miss: The psychology of rational thought.* Yale University Press. [RJS]

Sterelny, K. (2012) *The evolved apprentice*. MIT Press. [DCD]

Sterelny, K. (2013) The informational commonwealth. In: *Arguing about human nature: Contemporary debates*, ed. L S. M. Downes & E. Machery, pp. 274–88. Routledge, Taylor & Francis. [DCD]

Sternberg, R. J. (1997) What does it mean to be smart? *Educational Leadership* 54 (6):20–24. [RJS]

Sternberg, R. J., ed. (2002) *Why smart people can be so stupid*. Yale University Press. [RJS]

Sternberg, R. J. & Davidson, J. E. (1995) *The nature of insight*. MIT Press. [aBML]

Sternberg, R. J. & Jordan, J., eds. (2005) *Handbook of wisdom: Psychological perspectives.* Cambridge University Press. [RJS]

Stuhlmüller, A., Taylor, J. & Goodman, N. D. (2013) Learning stochastic inverses. Presented at the 2013 Neural Information Processing Systems conference, Lake Tahoe, NV, December 5–10, 2013. In: *Advances in Neural Information Processing Systems 26 (NIPS 2013)*, ed. C J. C. Burges, L. Bottou, M. Welling, Z. Ghagramani & K. Q. Weinberger, pp. 3048–56. Neural Information Processing Systems Foundation. [aBML]

Sukhbaatar, S., Szlam, A., Weston, J. & Fergus, R. (2015) End-to-end memory networks. Presented at the 2015 Neural Information Processing Systems conference, Montreal, QC, Canada, December 7–12, 2015. In: *Advances in neural information processing systems 28 (NIPS 2015)*, ed. C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama & R. Garnett [oral presentation]. Neural Information Processing Systems Foundation. [arBML]

Sun, R. (2016) *Anatomy of the mind*. Oxford University Press. [CDG]

Super, C. M. & Harkness, S. (2002) Culture structures the environment for development. *Human Development* 45(4):270–74. [JMC]

Sutton, R. S. (1990) Integrated architectures for learning, planning, and reacting based on approximating dynamic programming. In: *Proceedings of the 7th International Workshop on Machine Learning (ICML), Austin, TX*, pp. 216–24. International Machine Learning Society. [aBML]

Svedholm, A. M. & Lindeman, M. (2013) Healing, mental energy in the physics classroom: Energy conceptions and trust in complementary and alternative medicine in grade 10–12 students. *Science & Education* 22(3):677–94. [EJL]

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. & Rabinovich, A. (2014) Going deeper with convolutions. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, June 7–12, 2015*, pp. 1–9. IEEE. [aBML]

Tan, L. H., Spinks, J. A., Eden, G. F., Perfetti, C. A. & Siok, W. T. (2005) Reading depends on writing, in Chinese. *Proceedings of the National Academy of Sciences of the United States of America* 102(24):8781–85. [ED]

Tauber, S. & Steyvers, M. (2011) Using inverse planning and theory of mind for social goal inference. In: *Proceedings of the 33rd Annual Conference of the Cognitive Science Society, Boston, MA, July 20–23, 2011*, pp. 2480–85. Cognitive Science Society. [aBML]

Taylor, E. G. & Ahn, W.-K. (2012) Causal imprinting in causal structure learning. *Cognitive Psychology* 65:381–413. [EJL]

Téglás, E., Vul, E., Girotto, V., Gonzalez, M., Tenenbaum, J. B. & Bonatti, L. L. (2011) Pure reasoning in 12-month-old infants as probabilistic inference. *Science* 332(6033):1054–59. [aBML]

Tenenbaum, J. B., Kemp, C., Griffiths, T. L. & Goodman, N. D. (2011) How to grow a mind: Statistics, structure, and abstraction. *Science* 331(6022):1279–85. [aBML]

Thomaz, A. L. & Cakmak, M. (2013) Active social learning in humans and robots. In: *Social learning theory: Phylogenetic considerations across animal, plant, and microbial taxa*, ed. K. B. Clark, pp. 113–28. Nova Science. ISBN 978-1-62618-268-4. [KBC]

Thorwart, A. & Livesey E. J. (2016) Three ways that non-associative knowledge may affect associative learning processes. *Frontiers in Psychology* 7:2024. doi: 10.3389/fpsyg.2016.02024. [EJL]

Thurstone, L. L. (1919) The learning curve equation. *Psychological Monographs* 26 (3):2–33. [LRC]

Tian, Y. & Zhu, Y. (2016) Better computer Go player with neural network and long-term prediction. Presented at the International Conference on Learning Representations (ICLR), San Juan, Puerto Rico, May 2–4, 2016. *arXiv preprint 1511.06410*. Available at: https://arxiv.org/abs/1511.06410. [aBML]

Todd, P. M. & Gigerenzer, G. (2007) Environments that make us smart: Ecological rationality. *Current Directions in Psychological Science* 16(3):167–71. doi:10.1111/j.1467-8721.2007.00497.x. [PMP]

Tomai, E. & Forbus, K. (2008) Using qualitative reasoning for the attribution of moral responsibility. In: *Proceedings of the 30th Annual Conference of the Cognitive Science Society, Washington, DC, July 23–26, 2008*. Cognitive Science Society. [KDF]

Tomasello, M. (1999) *The cultural origins of human cognition*. Harvard University Press. [MHT]

Tomasello, M. (2010) *Origins of human communication*. MIT Press. [aBML]

Tompson, J. J., Jain, A., LeCun, Y. & Bregler, C. (2014). Joint training of a convolutional network and a graphical model for human pose estimation. Presented at the 28th Annual Conference on Neural Information Processing Systems, Montreal, Canada. In: *Advances in Neural Information Processing Systems* 27 (NIPS 2014), ed. Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence & K. Q. Weinberger, pp. 1799–807. Neural Information Processing Systems Foundation. [rBML]

Torralba, A., Murphy, K. P. & Freeman, W. T. (2007) Sharing visual features for multiclass and multiview object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29(5):854–69. [aBML]

Toshev, A. & Szegedy, C. (2014). Deeppose: Human pose estimation via deep neural networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH*, pp. 1653–60. IEEE. [rBML]

Tremoulet, P. D. & Feldman, J. (2000) Perception of animacy from the motion of a single object. *Perception* 29:943–51. [aBML]

Trettenbrein, P. C. (2016) The demise of the synapse as the locus of memory: A looming paradigm shift? *Frontiers in Systems Neuroscience* 10:88. [DG]

Tsividis, P., Gershman, S. J., Tenenbaum, J. B. & Schulz, L. (2013) Information selection in noisy environments with large action spaces. In: *Proceedings of the 36th Annual Conference of the Cognitive Science Society, Austin, TX*, pp. 1622–27. Cognitive Science Society. [aBML]

Tsividis, P., Tenenbaum, J. B. & Schulz, L. E. (2015) Constraints on hypothesis selection in causal learning. *Proceedings of the 37th Annual Conference of the Cognitive Sciences, Pasadena, CA, July 23–25, 2015*, pp. 2434–439. Cognitive Science Society. [aBML]

Tsividis, P. A., Pouncy, T., Xu, J. L., Tenenbaum, J. B. & Gershman, S. J. (2017) Human learning in Atari. In: *Proceedings of the Association for the Advancement of Artificial Intelligence (AAAI) Spring Symposium on Science of Intelligence: Computational Principles of Natural and Artificial Intelligence, Stanford University, Palo Alto, CA, March 25–27, 2017*. AAAI Press. [MHT, rBML]

Turing, A. M. (1950) Computing machine and intelligence. *Mind* 59:433–60. Available at: http://mind.oxfordjournals.org/content/LIX/236/433. [aBML]

Turovsky, B. (2016) Found in translation: More accurate, fluent sentences in Google Translate. Available at: https://blog.google/products/translate/found-translation-more-accurate-fluent-sentences-google-translate/. [DEM]

Tversky, B. & Hemenway, K. (1984) Objects, parts, and categories. *Journal of Experimental Psychology: General* 113(2):169–91. [aBML]

Ullman, S., Harari, D. & Dorfman, N. (2012a) From simple innate biases to complex visual concepts. *Proceedings of the National Academy of Sciences of the United States of America* 109(44):18215–20. [aBML]

Ullman, T. D., Baker, C. L., Macindoe, O., Evans, O., Goodman, N. D. & Tenenbaum, J. B. (2009). Help or hinder: Bayesian models of social goal inference. Presented at the 2009 Annual Conference on Neural Information Systems Processing, Vancouver, BC, Canada, December 7–10, 2009. In: *Advances in Neural Information Processing Systems 22 (NIPS 2009)*, ed. Y. Bengio, D. Schuumans, J. D. Lafferty, C. K. I. Williams & A. Culotta. Neural Information Processing Systems Foundation. [rBML]

Ullman, T. D., Goodman, N. D. & Tenenbaum, J. B. (2012b) Theory learning as stochastic search in the language of thought. *Cognitive Development* 27(4):455–80. [aBML]

U.S. Postal Service Historian (2016) Pieces of mail handled, number of post offices, income, and expenses since 1789. Available at: https://about.usps.com/who-we-are/postal-history/pieces-of-mail-since-1789.htm. [DG]

van den Hengel, A., Russell, C., Dick, A., Bastian, J., Pooley, D., Fleming, L. & Agapitol, L. (2015) Part-based modelling of compound scenes from images. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, June 7–12, 2015*, pp. 878–86. IEEE. [aBML]

van den Oord, A., Kalchbrenner, N. & Kavukcuoglu, K. (2016). Pixel recurrent neural networks. Presented at the 33rd International Conference on Machine Learning, New York, NY. *Proceedings of Machine Learning Research* 48:1747–56. [MB]

van Hasselt, H., Guez, A. & Silver, D. (2016) Deep learning with double Q-learning. In: *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence and the Twenty-Eighth Innovative Applications of Artificial Intelligence Conference on Artificial Intelligence, Phoenix, AZ*. AAAI Press. [aBML]

Varlet, M., Marin, L., Capdevielle, D., Del-Monte, J., Schmidt, R.C., Salesse, R.N., Boulenger, J.-P., Bardy, B. & Raffard, S. (2014). Difficulty leading interpersonal coordination: Towards an embodied signature of social anxiety disorder. *Frontiers in Behavioral Neuroscience* 8:1–9. [LM]

Vinyals, O., Blundell, C., Lillicrap, T. & Wierstra, D. (2016) Matching networks for one shot learning. Vinyals, O., Blundell, C., Lillicrap, T. Kavukcuoglu, K. & Wierstra, D. (2016). Matching networks for one shot learning. Presented at the 2016 Neural Information Processing Systems conference, Barcelona, Spain, December 5–10, 2016. In: *Advances in Neural Information Processing Systems 29 (NIPS 2016)*, ed. D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon & R. Garnett, pp. 3630–38. Neural Information Processing Systems Foundation. [arBML, MB, SSH]

Vinyals, O., Toshev, A., Bengio, S. & Erhan, D. (2014) Show and tell: A neural image caption generator. *arXiv preprint 1411.4555*. Available at: https://arxiv.org/abs/1411.4555. [aBML]

Viviani, P. & Stucchi, N. (1992). Biological movements look uniform: evidence of motor-perceptual interactions. *Journal of Experimental Psychology: Human, Perception & Performance* 18(3):603–23. [LM]

Vollmer, A-L., Mühlig, M., Steil, J. J., Pitsch, K., Fritsch, J., Rohlfing, K. & Wrede, B. (2014) Robots show us how to teach them: Feedback from robots shapes tutoring behavior during action learning, *PLoS One* 9(3):e91349. [P-YO]

Vul, E., Goodman, N., Griffiths, T. L. & Tenenbaum, J. B. (2014) One and done? Optimal decisions from very few samples. *Cognitive Science* 38(4):599–637. [aBML]

Vygotsky, L. S. (1978) Interaction between learning and development. In: *Mind in society: The development of higher psychological processes*, ed. M. Cole, V. John-Steiner, S. Scribner & E. Souberman, pp. 79–91. Harvard University Press. [RPK]

Wallach, W., Franklin, S. & Allen C. (2010) A conceptual and computational model of moral decision making in human and artificial agents. *Topics in Cognitive Science* 2:454–85. [KBC]

Wang, J. X., Kurth-Nelson, Z., Tirumala, D., Soyer, H., Leibo, J. Z., Munos, R., Blundell, C., Kumaran, D. & Botvinick, M. (2017). Learning to reinforcement learn. In: Presented at the 39th Annual Meeting of the Cognitive Science Society, London, July 26–29, 2017. *arXiv preprint 1611.05763*. Available at: https://arxiv.org/abs/1611.05763. [MB]

Wang, Z., Schaul, T., Hessel, M., Hasselt, H. van, Lanctot, M. & de Freitas, N. (2016) Dueling network architectures for deep reinforcement learning. *arXiv preprint 1511.06581*. Available at: http://arxiv.org/abs/1511.06581. [aBML]

Ward, T. B. (1994) Structured imagination: The role of category structure in exemplar generation. *Cognitive Psychology* 27:1–40. [aBML]

Watkins, C. J. & Dayan, P. (1992) Q-learning. *Machine Learning* 8:279–92. [aBML]

Weigmann, K. (2006) Robots emulating children. *EMBO Reports* 7(5):474–76. [KBC]

Weizenbaum, J. (1966) ELIZA—A computer program for the study of natural language communication between man and machine. *Communications of the ACM* 9(1):36–45. [RJS]

Wellman, H. M. & Gelman, S. A. (1992) Cognitive development: Foundational theories of core domains. *Annual Review of Psychology* 43:337–75. [arBML]

Wellman, H. M. & Gelman, S. A. (1998). Knowledge acquisition in foundational domains. In: *Handbook of child psychology: Vol. 2. Cognition, perception, and language development*, 5th ed., series ed. W. Damon, vol. ed. D. Kuhn & R. S. Siegler, pp. 523–73. Wiley. [arBML]

Weng, J., McClelland, J., Pentland, A., Sporns, O., Stockman, I., Sur, M. & Thelen, E. (2001) Autonomous mental development by robots and animals. *Science* 291 (5504):599–600. [GB]

Wermter, S., Palm, G., Weber, C. & Elshaw, M. (2005) Towards biomimetic neural learning for intelligent robots. In: *Biomimetic neural learning for intelligent robots*, ed. S. Wermter, G. Palm & M. Elshaw, pp. 1–18. Springer. [SW]

Weston, J., Bordes, A., Chopra, S., Rush, A. M., van Merriënboer, B., Joulin, A. & Mikolov, T. (2015a) Towards AI-complete question answering: A set of prerequisite toy tasks. *arXiv preprint 1502.05698*. Available at: https://arxiv.org/pdf/1502.05698.pdf. [SSH]

Weston, J., Chopra, S. & Bordes, A. (2015b) Memory networks. Presented at the International Conference on Learning Representations, San Diego, CA, May 7–9, 2015. arXiv:1410.3916. Available at: https://arxiv.org/abs/1410.3916. [arBML]

Williams, J. J. & Lombrozo, T. (2010) The role of explanation in discovery and generalization: Evidence from category learning. *Cognitive Science* 34(5):776–806. [aBML]

Wills, T. J., Cacucci, F., Burgess, N. & O'Keefe, J. (2010). Development of the hippocampal cognitive map in preweanling rats. *Science* 328(5985):1573–76. [ESS]

Winograd, T. (1972) Understanding natural language. *Cognitive Psychology* 3:1–191. [aBML, RJS]

Winston, P. H. (1975) Learning structural descriptions from examples. In: *The psychology of computer vision*, pp.157–210. McGraw-Hill. [aBML]

Wiskott, L. (2006). How does our visual system achieve shift and size invariance? In: *23 Problems in systems neuroscience*, ed. J. L. Van Hemmen & T. J. Sejnowski, pp. 322–40. Oxford University Press. [DG]

Wiskott, L. & von der Malsburg, C. (1996) Face recognition by dynamic link matching. In: *Lateral interactions in the cortex: structure and function*, ed. J. Sirosh, R. Miikkulainen and Y. Choe, ch 11. The UTCS Neural Networks Research Group. [DG]

Wolfram, S. (2002) *A new kind of science.* Wolfram Media. ISBN 1-57955-008-8. [KBC]

Wolpert, D. M., Miall, R. C. & Kawato, M. (1998) Internal models in the cerebellum. *Trends in Cognitive Science* 2(9):338–47. [GB]

Xu, F. & Tenenbaum, J. B. (2007) Word learning as Bayesian inference. *Psychological Review* 114(2):245–72. [aBML]

Xu, K., Ba, J., Kiros, R., Cho, K., Courville, A., Salakhutdinov, R., Zemel, R. & Bengio, Y. (2015) Show, attend and tell: Neural image caption generation with visual attention. Presented at the 2015 International Conference on Machine Learning. *Proceedings of Machine Learning Research* 37:2048–57. [arBML]

Yamada, Y., Mori, H. & Kuniyoshi, Y. (2010) A fetus and infant developmental scenario: Self-organization of goal-directed behaviors based on sensory constraints. In: *Proceedings of the 10th International Conference on Epigenetic Robotics, Õrenäs Slott, Sweden*, pp. 145–52. Lund University Cognitive Studies. [P-YO]

Yamins, D. L. K., Hong, H., Cadieu, C. F., Solomon, E. A., Seibert, D. & DiCarlo, J. J. (2014) Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proceedings of the National Academy of Sciences of the United States of America* 111(23):8619–24. [aBML, NK]

Yildirim, I., Kulkarni, T. D., Freiwald, W. A. & Tenenbaum, J. (2015) Efficient analysis-by-synthesis in vision: A computational framework, behavioral tests, and comparison with neural representations. In: *Proceedings of the 37th Annual Conference of the Cognitive Science Society, Pasadena, CA, July 22–25, 2015.* Cognitive Science Society. Available at: https://mindmodeling.org/cogsci2015/papers/0471/index.html. [aBML, NK]

Yosinski, J., Clune, J., Bengio, Y. & Lipson, H. (2014) How transferable are features in deep neural networks? Presented at the 2014 Neural Information Processing Systems conference, Montreal, QC, Canada. In: *Advances in neural information processing systems 27 (NIPS 2014)*, ed. Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence & K. Q. Weinberger [oral presentation]. Neural Information Processing Systems Foundation. [aBML]

Youyou, W., Kosinski, M. & Stillwell, D. (2015) Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences of the United States of America* 112(4):1036–40. [KBC]

Zeiler, M. D. & Fergus, R. (2014) Visualizing and understanding convolutional networks. In: *Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part I*, ed. D. Fleet, T. Pajdla, B. Schiele & T. Tuytelaars, pp. 818–33. Springer. [aBML]

Zentall, T. R. (2013) Observational learning in animals. In: *Social learning theory: Phylogenetic considerations across animal, plant, and microbial taxa*, ed. K. B. Clark, pp. 3–33. Nova Science. ISBN 978-1-62618-268-4. [KBC]

Zhou, H., Friedman, H. S. & von der Heydt, R. (2000) Coding of border ownership in monkey visual cortex. *The Journal of Neuroscience* 20:6594–611. [DGe]