The Measure of All Minds

Evaluating Natural and Artificial Intelligence

José Hernández-Orallo

jorallo@dsic.upv.es

Cambridge University Press 2017

The quintessence of intelligence is one of the big questions still beyond our understanding. In the past, science has unravelled many other previously puzzling questions through measurement, a fundamental tool for the identification, comparison and classification of natural phenomena. Not surprisingly, a very significant portion of our still scant knowledge about what intelligence is – and what it is not – comes from this measurement effort. For more than a century, psychometrics, comparative psychology and other disciplines have developed a rich collection of measurement instruments for quantifying various behavioural properties in the animal kingdom, prominently placing humans as a yardstick.

Beyond the enormous landscape of behaviours in the animal kingdom, there is yet another gigantic space to be explored: the machine kingdom. A plethora of new types of 'creatures' is emerging: robots, animats, chatbots, digital assistants, social bots, automated avatars and artificial life forms, to name a few, including hybrids and collectives, such as machine-enhanced humans, cyborgs, artificial swarms, human computation systems and crowd computing platforms. These systems display behaviours and capabilities as peculiar as their developers and constituents can contrive. Universal psychometrics presents itself as a new area dealing with the measurement of behavioural features in the machine kingdom, which comprises any interactive system, biological, artificial or hybrid, individual or collective.

The focus on an enlarged set of subjects generates plenty of new questions and opportunities. Are IQ tests valid for arbitrary machines? Can we devise universal cognitive tests? Can we have a formal definition of intelligence solely based on computational principles? Can the structure of cognitive abilities and empirical latent factors, including the dominant *g* factor, be extrapolated beyond biological creatures? Can this be studied theoretically? How should artificial personalities be measured? Do we need intelligence to evaluate intelligence universally? The classical paradigms used to evaluate natural and artificial systems have not been able to answer (or even formulate) these questions precisely. Also, customary evaluation tools are gamed by these new kinds of systems.

Recently, however, there has been a significant progress in a principled approach to the evaluation of behaviour based on information theory and computation. The anthropocentric stance is replaced by a universal perspective where life forms are considered as particular cases. Classical tools in human psychometrics, comparative psychology and animal cognition are not jettisoned but rethought for a wider landscape and substantiated on algorithmic grounds.

This book provides a comprehensive account of the concepts, terminology, theory and tools that should compose a unified framework for the universal evaluation of behavioural features. The exposition does not avoid some notions that are less consolidated, such as the arrangement of the space of abilities, the evaluation of personality or the process of ability development. The ideas that do not work are openly criticised, to aid the understanding of the many scattered scientific contributions that have recently appeared in different areas. In fact, some of these theories only make real sense – or no sense at all – when they are put together.

Many of the current conundrums in the evaluation of natural intelligence derive from the empirical evaluation of 'populations' (human groups, age ranges, species, etc.). The consideration of any conceivable behaviour (natural or artificial) and any imaginable 'machine population' provides a falsifiability criterion for any general claim, theory or test about behavioural features. The machine kingdom also brings a myriad of subjects to evaluate, with fewer experimentation constraints than those posed by humans and other animals. The theoretical underpinning on computation and information theory leads to several key formalisations, such as the concepts of task difficulty and policy-general intelligence. These new grounds illuminate blatant questions such as what human intelligence tests really measure.

Artificial intelligence can also benefit from the distinction between task-oriented evaluation and feature-oriented evaluation, jointly with a less anthropocentric methodology for the development and assessment of general-purpose agents. If properly overhauled, many tools from psychometrics can enter the scene of artificial intelligence evaluation, such as item response theory and adaptive testing. Similarly, the experience in the design of interfaces from animal evaluation can be crucial beyond natural intelligence.

Psychometrics, comparative psychology and artificial intelligence evaluation usually speak different languages. A great effort has been made to render this book accessible and valuable for researchers and students in all these areas and, extensively, to any interested reader outside these disciplines. As a result of the integration of different areas, some paradigms will be challenged and some hypotheses will be refuted. The outcome for the future is an integration of well-founded principles for the evaluation of behaviour in humans, non-human animals and all other machines.

References

Abbeel, P. and Ng, A. Y. (2004). Apprenticeship learning via inverse reinforcement learning. In *International Conference on Machine Learning*, pages 1–8. ACM International Conference Proceeding Series 69. ACM

Abramson, J. Z., Hernández-Lloreda, V., Call, J., and Colmenares, F. (2011). Relative quantity judgments in South American sea lions (*Otaria flavescens*). *Animal Cognition*, 14(5):695–706.

Adams, H. (1904). A law of acceleration. Chapter 34 *in (1905) "The education of Henry Adams"*. Houghton Mifflin.

Adams, S. S., Arel, I., Bach, J., Coop, R., Furlan, R., Goertzel, B., Hall, J. S., Samsonovich, A., Scheutz, M., Schlesinger, M., Shapiro, S. C., and Sowa, J. (2012). Mapping the landscape of human-level artificial general intelligence. *AI Magazine*, 33(1):25–42.

Adams, S. S., Banavar, G., and Campbell, M. (2016). I-athlon: Towards a multi-dimensional Turing test. *AI Magazine*, 37(1):78–84.

AERA, APA & NCME (1999). *Standards for educational and psychological testing*. American Educational Research Association.

AFADA v. Argentina (2015). Juzgado Nacional de Primera Instancia en lo Contencioso Administrativo Federal, Argentinian Court 34, 21-oct-2015, MJ-JU-M-95342-AR.

Agrillo, C. and Beran, M. J. (2013). Number without language: Comparative psychology and the evolution of numerical cognition. *Frontiers in Psychology*, 4.

Airiau, S. (2013). Cooperative games and multiagent systems. *The Knowledge Engineering Review*, 28(4):381–424.

Albert, D. and Held, T. (1999). Component-based knowledge spaces in problem solving and inductive reasoning. In D. Albert and J. Lukas (Eds.) *Knowledge spaces: Theories, empirical research, and applications*, pages 15–40. Psychology Press.

Alberts, J. R. and Ronca, A. E. (2012). The experience of being born: A natural context for learning to suckle. *International Journal of Pediatrics*, 2012:ID 129328, 11 pages.

Alcalá, J., Fernández, A., Luengo, J., Derrac, J., García, S., Sánchez, L., and Herrera, F. (2010). Keel data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework. *Journal of Multiple-Valued Logic and Soft Computing*, 17:255–287.

Alliger, G. M. (1988). Do zero correlations really exist among measures of different intellectual abilities? *Educational and Psychological Measurement*, 48(2):275–280.

Alur, R., Bodik, R., Juniwal, G., Martin, M. M. K., Raghothaman, M., Seshia, S. A., Singh, R., Solar-Lezama, A., Torlak, E., and Udupa, A. (2013). Syntax-guided synthesis. In B. Jobstmann and S. Ray (Eds.) *Formal Methods in Computer-Aided Design (FMCAD), 2013*, pages 1–17. IEEE.

Alvarado, N., Adams, S. S., Burbeck, S., and Latta, C. (2002). Beyond the Turing test: Performance metrics for evaluating a computer simulation of the human mind. In *International Conference on Development and Learning, 2002*, pages 147–152. IEEE.

Ambridge, B. and Lieven, E. V. (2011). *Child language acquisition: Contrasting theoretical approaches*. Cambridge University Press.

Amigoni, F., Bastianelli, E., Berghofer, J., Bonarini, A., Fontana, G., Hochgeschwender, N., Iocchi, L., Kraetzschmar, G., Lima, P., Matteucci, M., Miraldo, P., Nardi, D., and Schiaffonati, V. (2015). Competitions for benchmarking: Task and functionality scoring complete performance assessment. *IEEE Robotics & Automation Magazine*, 22(3):53–61. Benchmarks available at http://rockinrobotchallenge.eu.

Anderson, B. (1993). Evidence from the rat for a general factor that underlies cognitive performance and that relates to brain size: Intelligence? *Neuroscience Letters*, 153(1):98–102.

Anderson, B. (2000). The g factor in non-human animals. In Bock, G. R., Goode, J. A., and Webb, K., editors, *The nature of intelligence*, volume 233, pages 79–95. John Wiley.

Anderson, C. and Franks, N. R. (2003). Teamwork in animals, robots, and humans. *Advances in the Study of Behavior*, 33:1–48.

Anderson, J., Baltes, J., and Cheng, C. T. (2011). Robotics competitions as benchmarks for AI research. *The Knowledge Engineering Review*, 26(1):11–17.

Anderson, J. and Lebiere, C. (2003). The Newell Test for a theory of cognition. *Behavioral and Brain Sciences*, 26(5):587–601.

Anderson, M. L. (2004). A flexible approach to quantifying various dimensions of environmental complexity. In Messina, E. R. and Meystel, A. M., editors, *Measuring the performance and intelligence of systems: Proceedings of the 2004 PerMIS Workshop, August 24–26, 2004*, pages 192–197. NIST.

Antol, S., Agrawal, A., Lu, J., Mitchell, M., Batra, D., Zitnick, C. L., and Parikh, D. (2015). VQA: Visual question answering. CoRR arXiv preprint:1505.00468.

Antunes, L., Fortnow, L., Van Melkebeek, D., and Vinodchandran, N. V. (2006). Computational depth: Concept and applications. *Theoretical Computer Science*, 354(3):391–404.

Applewhite, P. (1975). Learning in bacteria, fungi, and plants. *Invertebrate Learning*, 3:179–186.

Arel, I. and Livingston, S. (2009). Beyond the Turing test. *Computer*, 42(3):90–91.

Arel, I., Rose, D. C., and Karnowski, T. P. (2010). Deep machine learning – a new frontier in artificial intelligence research. *Computational Intelligence Magazine*, 5(4):13–18.

Armstrong, S. and Sotala, K. (2015). How we're predicting AI–or failing to. In J. Romportl, E. Zackova and J. Kelemen (Eds.) *Beyond artificial intelligence*, pages 11–29. Springer.

Arnold, P. and Heiron, K. (2002). Tactile memory of deaf-blind adults on four tasks. *Scandinavian Journal of Psychology*, 43(1):73–79.

Arora, S. and Barak, B. (2009). *Computational complexity: A modern approach*. Cambridge University Press.

Arrabales, R., Ledezma, A., and Sanchis, A. (2010). ConsScale: A pragmatic scale for measuring the level of consciousness in artificial agents. *Journal of Consciousness Studies*, 17(3–4):131–164.

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

Ashby, W. R. (1956). *An introduction to cybernetics*. Chapman and Hall.

Ashcraft, M. H., Donley, R. D., Halas, M. A., and Vakali, M. (1992). Working memory, automaticity, and problem difficulty. *Advances in Psychology*, 91:301–329.

Ashlock, D. and Schonfeld, J. (2010). Evolution for automatic assessment of the difficulty of Sokoban boards. In *IEEE Congress on Evolutionary Computation (CEC)*, pages 1–8. IEEE.

Asimov, I. (1950). *I, robot*. Gnome Press.

ASPCR (1999). American Society for the Prevention of Cruelty to Robots. http://www.aspcr.com/.

Ay, N., Mueller, M., and Szkola, A. (2008). Effective complexity and its relation to logical depth. CoRR ArXiv preprint:0810.5663.

Aziz, H., Brill, M., Fischer, F., Harrenstein, P., Lang, J., and Seedig, H. G. (2015). Possible and necessary winners of partial tournaments. *Journal of Artificial Intelligence Research*, 54:493–534.

Bache, K. and Lichman, M. (2013). UCI machine learning repository. University of California, Irvine, School of Information and Computer Sciences. http://archive.ics.uci.edu/ml.

Bachrach, Y., Graepel, T., Kasneci, G., Kosinski, M., and Van Gael, J. (2012a). Crowd IQ: Aggregating opinions to boost performance. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, volume 1, pages 535–542.

Bachrach, Y., Kosinski, M., Graepel, T., Kohli, P., and Stillwell, D. (2012b). Personality and patterns of facebook usage. In *Proceedings of the 3rd Annual ACM Web Science Conference*, pages 24–32. ACM.

Bacon, F. (1718). *The essays, or councils, civil and moral*. H. Clark.

Badler, N., Allbeck, J., Zhao, L., and Byun, M. (2002). Representing and parameterizing agent behaviors. In *Proceedings of Computer Animation, 2002*, pages 133–143. IEEE.

Balch, T. (2000). Hierarchic social entropy: An information theoretic measure of robot group diversity. *Autonomous Robots*, 8(3):209–238.

Balch, T., Bellman, K., Cotsfatis, M., Davis, P., Davis, W. J., Fakory, R., Finkelstein, R., Grant, E., Hernández-Orallo, J., Joslyn, C., Reeker, L., Messina, E., Meystel, A., Murphy, R., Peterson, C., Phoha, S., Pouchard, L., Samad, T., Sanderson, A., Schultz, A., Stirling, W., Sukhatme, G., Wallace, S., Wild, A., Weng, J., and Whalen, T. (2000). Compendium of the minutes of plenary discussion. In Meystel, A. M. and Messina, E. R., editors, *Measuring the performance and intelligence of systems: Proceedings of the 2000 PerMIS Workshop*, pages 543–556. NIST Special Publication 970, Gaithersburg, MD.

Baldassarre, G. and Mirolli, M., editors (2013). *Intrinsically motivated learning in natural and artificial systems*. Springer.

Barenbaum, N. B. and Winter, D. G. (2008). History of modern personality theory and research. In *Handbook of personality: Theory and research*, pages 3–26. Guilford Press.

Barlow, J. B. and Dennis, A. R. (2014). Not as smart as we think: A study of collective intelligence in virtual groups. *Proceedings of Collective Intelligence*.

Barmpalias, G. and Dowe, D. L. (2012). Universality probability of a prefix-free machine. *Philosophical Transactions of the Royal Society, Series*, 370:3488– 3511.

Baroody, A. J. and Dowker, A. (2013). *The development of arithmetic concepts and skills: Constructive adaptive expertise*. Routledge.

Barrera, A., Cáceres, A., Weitzenfeld, A., and Ramírez-Amaya, V. (2011). Comparative experimental studies on spatial memory and learning in rats and robots. *Journal of Intelligent and Robotic Systems*, 63:361–397.

Barrett, G. V. and Depinet, R. L. (1991). A reconsideration of testing for competence rather than for intelligence. *American Psychologist*, 46(10):1012.

Barrett, P. (2004). The future of ability testing: What happened to the science of mental ability. *Measurement*, 2(1):33–54.

Barrett, S. and Stone, P. (2012). An analysis framework for ad hoc teamwork tasks. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, Volume 1, pages 357–364. International Foundation for Autonomous Agents and Multiagent Systems.

Bartholomew, D. J. (2004). *Measuring intelligence: Facts and fallacies*. Cambridge University Press.

Bates, J., et al. (1994). The role of emotion in believable agents. *Communications of the ACM*, 37(7):122–125.

Baum, E. B. (2004). *What is thought?* MIT Press.

Bayley, N. (1969). *Scales of infant development*. *Psychological Corporation*.

Bayley, N. (1993). *Bayley scales of infant development: Manual*. Psychological Corporation.

BBC (2003). Is txt mightier than the word? *BBC News*, 4 March. http://news.bbc.co.uk/go/pr/fr/-/2/hi/uk_news/2814235.stm.

BBC (2014a). Court in Argentina grants basic rights to orangutan. *BBC News*, 21 December. http://www.bbc.com/news/world-latin-america-30571577.

BBC (2014b). US chimpanzee Tommy 'has no human rights' – court. *BBC News*, 4 December. http://www.bbc.com/news/world-us-canada-30338231.

Beck, B. B. (1982). Chimpocentrism: Bias in cognitive ethology. *Journal of Human Evolution*, 11(1):3–17.

Beer, R. (2004). Autopoiesis and cognition in the game of life. *Artificial Life*, 10(3):309–326.

Beer, R. D. (2003). The dynamics of active categorical perception in an evolved model agent. *Adaptive Behavior*, 11(4):209–243.

Beer, R. D. (2015). Characterizing autopoiesis in the game of life. *Artificial Life*, 21(1):1–19.

Beer, R. D. and Williams, P. L. (2015). Information processing and dynamics in minimally cognitive agents. *Cognitive Science*, 39(1):1–38.

Bejar, I. I., Chaffin, R., and Embretson, S. (1991). *Cognitive and psychometric analysis of analogical problem solving*. Springer.

Bellemare, M. G., Naddaf, Y., Veness, J., and Bowling, M. (2013). The arcade learning environment: An evaluation platform for general agents. *Journal of Artificial Intelligence Research*, 47:253–279.

Ben-Jacob, E. (2008). Social behavior of bacteria: From physics to complex organization. *European Physical Journal B*, 65(3):315–322.

Ben-Jacob, E. (2009). Learning from bacteria about natural information processing. *Annals of the New York Academy of Sciences*, 1178(1):78–90.

Ben-Porath, Y. S. (2012). *Interpreting the MMPI-2-RF*. University of Minnesota Press.

Bender, D. (2015). Establishing a human baseline for the Winograd schema challenge. In M. Glass and J. H. Kim (Eds.), *Modern AI and Cognitive Science Conference, CEUR Workshop Proceedings*, 1353:39–45.

Bennett, C. H., Gács, P., Li, M., Vitányi, P. M., and Zurek, W. H. (1998). Information distance. *IEEE Transactions on Information Theory*, 44(4):1407–1423.

Benson-Amram, S., Dantzer, B., Stricker, G., Swanson, E. M., and Holekamp, K. E. (2016). Brain size predicts problem-solving ability in mammalian carnivores. *Proceedings of the National Academy of Sciences*, 113(9):2352–2357.

Bentham, J. (1780). *An introduction to the principles of morals and legislation*. T. Payne.

Bentley, J. G. W., Bishop, P. G., and Van Der Meulen, M. (2004). An empirical exploration of the difficulty function. In M. Heisel, P. Liggesmeyer, S. Wittman (Eds.) *Computer Safety, Reliability, and Security*, LNCS 3219, pages 60–71. Springer.

Berant, J., Srikumar, V., Chen, P.-C., Huang, B., Manning, C. D., Vander Linden, A., Harding, B., and Clark, P. (2014). Modeling biological processes for reading comprehension. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing, EMNLP*, pages 1499–1510. Association for Computational Linguistics.

Berg-Cross, G. (2007). Report on panel discussion on can the development of intelligent robots be benchmarked? Concepts and issues from epigenetic robotics. In Messina, E. R. and Madhavan, R., editors, *Proceedings of the 2007 Workshop on Performance Metrics for Intelligent Systems*, pages 167–181. NIST.

Berger, T. W., Song, D., Chan, R. H., Marmarelis, V. Z., LaCoss, J., Wills, J., Hampson, R. E., Deadwyler, S., Granacki, J. J., et al. (2012). A hippocampal cognitive prosthesis: Multi-input, multi-output nonlinear modeling and VLSI implementation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 20(2):198–211.

Bernasconi, A., Bodei, C., and Pagli, L. (2007). Knitting for fun: A recursive sweater. In *Fun with algorithms*, pages 53–65. Springer.

Bernstein, A., Klein, M., and Malone, T. W. (2012). Programming the global brain. *Communications of the ACM*, 55(5):41–43.

Besold, T. K., Hernández-Orallo, J., and Schmid, U. (2015). Can machine intelligence be measured in the same way as human intelligence? *Künstliche Intelligenz*, 29(3):291–297.

Besold, T. R. (2014). A note on chances and limitations of Psychometric AI. In *KI 2014: Advances in Artificial Intelligence*, pages 49–54. Springer.

Betjemann, R. S., Johnson, E. P., Barnard, H., Boada, R., Filley, C. M., Filipek, P. A., Willcutt, E. G., DeFries, J. C., and Pennington, B. F. (2010). Genetic covariation between brain volumes and IQ, reading performance, and processing speed. *Behavior genetics*, 40(2):135–145.

Beuls, K., Steels, L., et al. (2013). Agent-based models of strategies for the emergence and evolution of grammatical agreement. *PloS ONE*, 8(3):e58960.

Bibel, W. (1980). Intellektik statt KI – Ein ernstgemeinter Vorschlag. *Rundbrief der Fachgruppe Künstliche Intelligenz in der Gesellschaft für Informatik*, 22:15–16.

Biemans, N. (2012). 'Killjoys' challenge claims of clever animals. *Science*, 335:1036–1037.

Bien, Z., Bang, W. C., Kim, D. Y., and Han, J. S. (2002). Machine intelligence quotient: Its measurements and applications. *Fuzzy Sets and Systems*, 127(1):3–16.

Bien, Z., Kim, Y. T., and Yang, S. H. (1998). How to measure the machine intelligence quotient (MIQ): Two methods and applications. In *World Automation Congress (WAC)*. TSI Press.

Biever, C. (2011). Ultimate IQ: One test to rule them all. *New Scientist*, 211(2829):42–45.

Binet, A. (1909). *Les idées modernes sur les enfants*. Flammarion.

Binet, A. and Simon, T. (1905). New methods for the diagnosis of the intellectual level of subnormals. *L'année Psychologique*, 12:191–244.

Birch, T. H. (1993). Moral considerability and universal consideration. *Environmental Ethics*, 15(4):313–332.

Bird, C. D. and Emery, N. J. (2009). Rooks use stones to raise the water level to reach a floating worm. *Current Biology*, 19(16):1410–1414.

Birder, L., Kanai, A., Cruz, F., Moore, K., and Fry, C. (2010). Is the urothelium intelligent? *Neurourology and Urodynamics*, 29(4):598–602.

Bjorklund, D. F. and Pellegrini, A. D. (2002). *The origins of human nature: Evolutionary developmental psychology.* American Psychological Association.

Bloom, P. and German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition*, 77(1):B25–B31.

Blum, L. and Blum, M. (1975). Toward a mathematical theory of inductive inference. *Information and Control*, 28(2):125–155.

Blumenbach, J. (1790). *Beytrage zur Naturgeschichte*, 2 vols. Göttingen: Dieterich.

Boake, C. (2002). From the Binet–Simon to the Wechsler–Bellevue: Tracing the history of intelligence testing. *Journal of Clinical and Experimental Neuropsychology*, 24(3):383–405.

Bolhuis, J. J. and Everaert, M. (2013). *Birdsong, speech, and language: Exploring the evolution of mind and brain*. MIT Press.

Bolhuis, J. J., Tattersall, I., Chomsky, N., and Berwick, R. C. (2014). How could language have evolved? *PLOS Biology*, 12(8).

Bonabeau, E., Dorigo, M., and Theraulaz, G. (1999). *Swarm intelligence: From natural to artificial systems*. Oxford University Press.

Borges, J. L. (1939). *The total library*. In The total library: Non-fiction 1922–1986, translated by Eliot Weinberge, pages 214–216 Penguin Press, London, 2000. Pages 214–216. Translated by Eliot Weinberge.

Borges, J. L. (1957). *The book of imaginary beings*. Originally published in Portuguese as *Manual de zoología fantástica*, Fce-Breviarios; English version 2002, Random House.

Borges, J. L. (1964). *Other inquisitions (1937–1952), Translated by Ruth L. C. Simms*. University of Texas Press.

Boring, E. G. (1923). Intelligence as the tests test it. *New Republic*, pages 35–37.

Borsboom, D. (2005). *Measuring the mind: Conceptual issues in contemporary psychometrics*. Cambridge University Press.

Borsboom, D., Mellenbergh, G., and Van Heerden, J. (2004). The concept of validity. *Psychological Review*, 111(4):1061.

Bosse, T., Jonker, C. M., Schut, M. C., and Treur, J. (2006). Collective representational content for shared extended mind. *Cognitive Systems Research*, 7(2):151–174.

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.

Bostrom, N. and Sandberg, A. (2009). Cognitive enhancement: Methods, ethics, regulatory challenges. *Science and Engineering Ethics*, 15(3):311–341.

Bouchard, T. J. (2014). Genes, evolution and intelligence. *Behavior Genetics*, 44(6):549–577.

Bowerman, M. and Levinson, S. C. (2001). *Language acquisition and conceptual development*, volume 3. Cambridge University Press.

Bradley-Johnson, S. and Morgan, S. K. (2008). *Psychoeducational assessment of students who are visually impaired or blind: Infancy through high school*. Region 4 Education Service Center.

Brazdău, O. and Mihai, C. (2011). The consciousness quotient: A new predictor of the students' academic performance. *Procedia – Social and Behavioral Sciences*, 11:245–250.

Bringsjord, S. (2000). In light of artificial intelligence, the science of mental ability is either silly or pointless. *Psycoloquy*, 11, Article 43.

Bringsjord, S. (2011). Psychometric artificial intelligence. *Journal of Experimental & Theoretical Artificial Intelligence*, 23(3):271–277.

Bringsjord, S., Bello, P., and Ferrucci, D. (2003). Creativity, the Turing test, and the (better) Lovelace test. In Moor, J. H., editor, *The Turing test: The elusive standard of artificial intelligence*, pages 215–239. Springer.

Bringsjord, S. and Schimanski, B. (2003). What is artificial intelligence? Psychometric AI as an answer. In *International Joint Conference on Artificial Intelligence*, pages 887–893. Morgan Kaufmann.

Broad, C. D. (1926). *The philosophy of Francis Bacon*. Cambridge University Press.

Brockman, J. (2015). What do you think about machines that think?, Annual question. *Edge.org*. http://edge.org/annual-question/what-do-you-think-about-machines-that-think.

Brooks, R. A. (1990). Elephants don't play chess. *Robotics and Autonomous Systems*, 6(1):3–15.

Brooks, R. R. (2001). *Stigmergy – an intelligence metric for emergent distributed behaviors*. In *Workshop on Performance Metrics for Intelligent Systems*, pages 186–192. NIST.

Brown, J. I., Fishco, V. V., and Hanna, G. (1993). *Nelson-Denny reading test: Manual for scoring and interpretation, forms G & H*. Riverside.

Brundage, M. (2016). Modeling progress in AI. In *AAAI 2016 Workshop on AI, Ethics, and Society*.

Bryson, J. J. (2010). Robots should be slaves. In Wilks, Y., editor, *Close engagements with artificial companions: Key social, psychological, ethical and design issues*, pages 63–74. John Benjamins.

Budoff, M. (1987). The validity of learning potential assessment. In Lidz, C. S., editor, *Dynamic testing*, pages 173–195. Guilford Press.

Buhrmester, M., Kwang, T., and Gosling, S. D. (2011). Amazon's mechanical turk a new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 6(1):3–5.

Bürckert, H.-J., Müller, J., and Schupeta, A. (1991). Ratman and its relation to other multi-agent testbeds. Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, Research Report, RR-91-09.

Burghart, C. R. and Steinfeld, A. (2008). Human-robot interaction metrics and future directions. In *Workshop on Metrics for Human-Robot Interaction*, page 1. School of Computer Science, University of Hertforshire.

Burman, E. (2007). *Deconstructing developmental psychology*. Routledge.

Burns, E. (1998). *Test accommodations for students with disabilities.* ERIC.

Bursztein, E., Aigrain, J., Moscicki, A., and Mitchell, J. C. (2014). The end is nigh: Generic solving of text-based CAPTCHAs. In *Proceedings of the 8th USENIX Conference on Offensive Technologies*, pages 3–3. USENIX Association.

Burt, C. (1949). The structure of the mind: A review of the results of factor analysis. *British Journal of Educational Psychology*, 19(2–3):100–111, 176–199.

Buschkuehl, M. and Jaeggi, S. M. (2010). Improving intelligence: A literature review. *Swiss Medical Weekly*, 140(19–20):266–272.

Busoniu, L., Babuska, R., and De Schutter, B. (2008). A comprehensive survey of multiagent reinforcement learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 38(2):156–172.

Buss, D. M. (1999). *Evolutionary psychology: The new science of the mind*. Allyn and Bacon.

Butler, S. (1872). *Erewhon*. William Brendon.

Call, J. and Tomasello, M. (1998). The effect of humans on the cognitive development of apes. In Russon, A. E., Bard, K. A., and Parker, S. T., editors, *Reaching into thought: The minds of the great apes*, pages 371–403. Cambridge University Press.

Callahan, C. M. and Hertberg-Davis, H. L. (2012). *Fundamentals of gifted education: Considering multiple perspectives*. Routledge.

Calvo-Garzón, P. (2003). Nonclassical connectionism should enter the decathlon. *Behavioral and Brain Sciences*, 26(05):603–604.

Calvo-Garzón, P. and Keijzer, F. (2011). Plants: Adaptive behavior, root-brains, and minimal cognition. *Adaptive Behavior*, 19(3):155–171.

Campbell, M., Hoane, A. J., and Hsu, F. (2002). Deep Blue. *Artificial Intelligence*, 134(1–2):57–83.

Candland, D. K. (1995). *Feral children and clever animals: Reflections on human nature*. Oxford University Press.

Cangelosi, A. and Schlesinger, M. (2015). *Developmental robotics: From babies to robots*. MIT Press.

Cantlon, J. F. and Brannon, E. M. (2006). Shared system for ordering small and large numbers in monkeys and humans. *Psychological Science*, 17(5):401–406.

Carbonell, J. G. (1980). Towards a process model of human personality traits. *Artificial Intelligence*, 15(1):49–74.

Carlson, J. F., Geisinger, K. F., and Jonson, J. L. (2014). *The nineteenth mental measurements yearbook*. Buros Center for Testing. http://buros.org/mental-measurements-yearbook.

Carpenter, P. A., Just, M. A., and Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices test. *Psychological Review*, 97:404–431.

Carr, N. (2008). Is Google making us stupid? *Yearbook of the National Society for the Study of Education*, 107(2):89–94.

Carr, N. (2011). *The shallows: What the Internet is doing to our brains*. W. W. Norton.

Carroll, J. B. (1981). Ability and task difficulty in cognitive psychology. *Educational Researcher*, pages 11–21.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press.

Carroll, J. L. and Seppi, K. (2005). Task similarity measures for transfer in reinforcement learning task libraries. In *International Joint Conference on Neural Networks*, volume 2, pages 803–808. IEEE.

Carter, M. (2007). *Minds and computers: An introduction to the philosophy of artificial intelligence*. Edinburgh University Press.

Castelfranchi, C., Rosis, F. D., Falcone, R., and Pizzutilo, S. (1998). Personality traits and social attitudes in multiagent cooperation. *Applied Artificial Intelligence*, 12(7–8):649–675.

Cattell, R. B. (1946). *Description and measurement of personality*. World Book Company.

Cattell, R. B., Eber, H. W., and Tatsuoka, M. M. (1970). *Handbook for the sixteen personality factor questionnaire (16 PF): In clinical, educational, industrial, and research psychology, for use with all forms of the test*. Institute for Personality and Ability Testing.

Celli, F., Lepri, B., Biel, J.-I., Gatica-Perez, D., Riccardi, G., and Pianesi, F. (2014). The workshop on computational personality recognition 2014. In *Proceedings of the ACM International Conference on Multimedia*, pages 1245–1246. ACM.

Celli, F., Pianesi, F., Stillwell, D., and Kosinski, M. (2013). The workshop on computational personality recognition 2013. AAAI Technical Report WS-13-01.

Chaitin, G. J. (1975). A theory of program size formally identical to information theory. *Journal of the ACM*, 22(3):329–340.

Chaitin, G. J. (1977). Algorithmic information theory. *IBM Journal of Research and Development*, 21(4):350–359.

Chaitin, G. J. (1982). Gödel's theorem and information. *International Journal of Theoretical Physics*, 21(12):941–954.

Chaitin, G. J. (2002). On the intelligibility of the universe and the notions of simplicity, complexity and irreducibility. In *Grenzen und Grenzüberschreitungen, XIX. Deutscher Kongress für Philosophie, Bonn, 23.-27. September 2002*. ArXiv preprint: math/0210035.

Chalmers, D. (2010). The singularity: A philosophical analysis. *Journal of Consciousness Studies*, 17(9-10):7–65.

Chandrasekaran, B. (1990). What kind of information processing is intelligence? In *The foundation of artificial intelligence – a sourcebook*, pages 14–46. Cambridge University Press.

Chang, N. C.-L. (2008). *Constructing grammar: A computational model of the emergence of early constructions*. PhD thesis, University of California at Berkeley.

Chapouthier, G. and Kaplan, F. (2011). *L'homme, l'animal et la machine*. CNRS éditions.

Chapouthier, G., Nouët, J. C., et al. (1998). *The universal declaration of animal rights: Comments and intentions*. Ligue Française des Droits de l'Animal.

Chappell, J. and Sloman, A. (2007). Natural and artificial meta-configured altricial information-processing systems. *International Journal of Unconventional Computing*, 3(3):211–239.

Chater, N. (1999). The search for simplicity: A fundamental cognitive principle? *Quarterly Journal of Experimental Psychology: Section A*, 52(2):273–302.

Chater, N., Clark, A., Perfors, A., and Goldsmith, J. A. (2015). *Empiricism and language learnability*. Oxford University Press.

Chater, N. and Vitányi, P. (2003). Simplicity: A unifying principle in cognitive science? *Trends in Cognitive Sciences*, 7(1):19–22.

Chater, N. and Vitányi, P. (2007). 'Ideal learning' of natural language: Positive results about learning from positive evidence. *Journal of Mathematical Psychology*, 51(3):135–163.

Chedid, F. B. (2010). Sophistication and logical depth revisited. In *2010 IEEE/ACS International Conference on Computer Systems and Applications (AICCSA)*, pages 1–4. IEEE.

Chemero, A. (2011). *Radical embodied cognitive science*. MIT Press.

Chmait, N., Dowe, D. L., Green, D. G., and Li, Y.-F. (2015a). Observation, communication and intelligence in agent-based systems. In Bieger, J., Goertzel, B., and Potapov, A., editors, *Artificial General Intelligence – 8th International Conference, AGI 2015, Berlin, Germany, July 22–25, 2015, Proceedings*, pages 50–59. Springer.

Chmait, N., Dowe, D. L., Green, D. G., Li, Y.-F., and Insa-Cabrera, J. (2015b). Measuring universal intelligence in agent-based systems using the anytime intelligence test. Technical Report 2015/279, Faculty of Information Technology, Clayton, Monash University.

Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.

Chomsky, N. (1998). *Minimalist inquiries: The framework*. MIT Working Papers in Linguistics.

Christiansen, M. H. and Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, 31(5):489–509.

Chu, Z., Gianvecchio, S., Wang, H., and Jajodia, S. (2012). Detecting automation of Twitter accounts: Are you a human, bot, or cyborg? *IEEE Transactions on Dependable and Secure Computing*, 9(6):811–824.

Cianciolo, A. T. and Sternberg, R. J. (2008). *Intelligence: A brief history*. John Wiley.

Cilibrasi, R. and Vitányi, P. M. (2005). Clustering by compression. *IEEE Transactions on Information Theory*, 51(4):1523–1545.

Clark, A. (2004). *Natural-born cyborgs: Minds, technologies, and the future of human intelligence*. Oxford University Press.

Clark, A. and Chalmers, D. (1998). The extended mind. *Analysis*, 58(1):7–19.

Clark, P. (2015). Elementary school science and math tests as a driver for AI: Take the Aristo challenge! *Proceedings of the Twenty-Seventh Conference on Innovative Applications of Artificial Intelligence*, pages 4019–4021.

Clark, P. and Etzioni, O. (2016). My computer is an honor student – but how intelligent is it? Standardized tests as a measure of AI. *AI Magazine*, 37(1):5–12.

Cliff, D. and Miller, G. F. (1995). Co-evolution of pursuit and evasion II: Simulation methods and results. In Maes, P., Mataric, M., Meyer, J., Pollack, J., and Wilson, S., editors, *From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, pages 506–515. MIT Press.

Cobham, A. (1965). The intrinsic computational difficulty of functions. In Yehoshua Bar-Hillel (Ed.), *Proceedings of the 1964 Congress for Logic, Methodology, and the Philosophy of Science*, pages 24–30. Studies in Logic and the Foundations of Mathematics. North-Holland.

Cochran, W. G. (1977). *Sampling techniques*. John Wiley.

Cohen, A., Fiorello, C. A., and Farley, F. H. (2006). The cylindrical structure of the Wechsler Intelligence Scale for Children IV: A retest of the Guttman model of intelligence. *Intelligence*, 34(6):587–591.

Cohen, P. R. (2005). If not Turing's test, then what? *AI Magazine*, 26(4):61.

Cohen, P. R. and Howe, A. E. (1988). How evaluation guides AI research: The message still counts more than the medium. *AI Magazine*, 9(4):35.

Cohen, Y. (2013). Testing and cognitive enhancement. Technical report, National Institute for Testing and Evaluation, Jerusalem.

Comins, J. A. and Gentner, T. Q. (2010). Working memory for patterned sequences of auditory objects in a songbird. *Cognition*, 117(1):38–53.

Commons, M. L., Trudeau, E. J., Stein, S. A., Richards, F. A., and Krause, S. R. (1998). Hierarchical complexity of tasks shows the existence of developmental stages. *Developmental Review*, 18(3):237–278.

Conway, A. R. A. and Kovacs, K. (2015). New and emerging models of human intelligence. *Wiley Interdisciplinary Reviews: Cognitive Science*, 6(5):419–426.

Conway, J. (1970). The game of life. *Scientific American*, 223(4):4.

Cook, M. (2004). Universality in elementary cellular automata. *Complex Systems*, 15(1):1–40.

Cook, R. G., Katz, J. S., and Cavoto, B. R. (1997). Pigeon same–different concept learning with multiple stimulus classes. *Journal of Experimental Psychology: Animal Behavior Processes*, 23(4):417.

Copeland, B. J. (2004). *The essential Turing: Seminal writings in computing, logic, philosophy, artificial intelligence, and artificial life, plus the secrets of Enigma*. Clarendon Press.

Copeland, J. and Proudfoot, D. (2008). Turing's test: A philosophical and historical guide. In Epstein, R., Roberts, G., and Beber, G., editors, *Parsing the Turing test: Philosophical and methodological issues in the quest for the thinking computer*. Springer.

Corballis, M. C. (2011). *The recursive mind: The origins of human language, thought, and civilization*. Princeton University Press.

Costa, P. T. and McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual Differences*, 13(6):653–665.

Costa, P. T. and McCrae, R. R. (2005). *Personality in adulthood: A five-factor theory perspective, second edition*. Routledge.

Cox, C. (1926). *The early mental traits of 300 geniuses*. PhD thesis, Stanford University.

Cox, J. J., Reimann, F., Nicholas, A. K., Thornton, G., Roberts, E., Springell, K., Karbani, G., Jafri, H., Mannan, J., Raashid, Y., Al-Gazali, L., Hamamy, H., Valente, E. M., Gorman, S., Williams, R., McHale, D. P., Wood, J. N., Gribble, F. M., and Woods, C. G. (2006). An SCN9A channelopathy causes congenital inability to experience pain. *Nature*, 444(7121):894–898.

Crinella, F. M. and Yu, J. (1995). Brain mechanisms in problem solving and intelligence: A replication and extension. *Intelligence*, 21(2):225–246.

DARPA (2015). Defense advanced research projects agency robotics challenge. http://www.theroboticschallenge.org/.

Darwin, C. (1871). *The descent of man, and selection in relation to sex*. John Murray.

Dautenhahn, K. (2007). Socially intelligent robots: Dimensions of human–robot interaction. *Philosophical Transactions of the Royal Society, Series*, 362(1480):679–704.

Davis, E. (2014). The limitations of standardized science tests as benchmarks for artificial intelligence research: Position paper. CoRR arXiv preprint:1411.1629.

Davis, E. (2016). How to write science questions that are easy for people and hard for computers. *AI Magazine*, 37(1):13–22.

Davydov, V. (1986). *Problemy razvivayushego obucheniya (issues in developing learning)*. Pedagogika.

Dawkins, M. S. (2008). The science of animal suffering. *Ethology*, 114(10):937–945.

Dawkins, M. S. (2012). *Animal suffering: The science of animal welfare*. Springer Science and Business Media.

Dawkins, R. (1976). *The selfish gene*. Oxford University Press.

Dawkins, R. (2000). *Unweaving the rainbow: Science, delusion and the appetite for wonder*. Houghton Mifflin Harcourt.

Dawkins, R. and Krebs, J. R. (1978). Animal signals: Information or manipulation. *Behavioural Ecology: An Evolutionary Approach*, 2:282–309.

De Ayala, R. J. (2009). *Theory and practice of item response theory*. Guilford Press.

De Condorcet, M. (1785). *Essais sur l'application de l'analyse à la probabilité des décisions rendues a la pluralité des voix*. Paris.

De Waal, F. (2013). *The bonobo and the atheist: In search of humanism among the primates*. W. W. Norton.

Deacon, R. M., Bannerman, D. M., Kirby, B. P., Croucher, A., and Rawlins, J. N. P. (2002). Effects of cytotoxic hippocampal lesions in mice on a cognitive test battery. *Behavioural Brain Research*, 133(1):57–68.

Deadwyler, S. A., Berger, T. W., Sweatt, A. J., Song, D., Chan, R. H., Opris, I., Gerhardt, G. A., Marmarelis, V. Z., and Hampson, R. E. (2013). Donor/recipient enhancement of memory in rat hippocampus. *Frontiers in Systems Neuroscience*, 7:Article 120.

Deary, I. J. (2000a). *Looking down on human intelligence: From psychometrics to the brain.* Oxford University Press.

Deary, I. J. (2000b). Simple information processing and intelligence. In Sternberg, R., editor, *Handbook of intelligence*, pages 267–284. Cambridge University Press.

Deary, I. J. and Caryl, P. G. (1997). Neuroscience and human intelligence differences. *Trends in Neurosciences*, 20(8):365–371.

Deary, I. J., Der, G., and Ford, G. (2001). Reaction times and intelligence differences: A population-based cohort study. *Intelligence*, 29(5):389–399.

Deary, I. J., Egan, V., Gibson, G. J., Austin, E. J., Brand, C. R., and Kellaghan, T. (1996). Intelligence and the differentiation hypothesis. *Intelligence*, 23(2):105– 132.

Decker, K. S., Durfee, E. H., and Lesser, V. R. (1989). Evaluating research in cooperative distributed problem solving. *Distributed Artificial Intelligence*, 2:487–519.

Dehaene, S., Izard, V., Pica, P., and Spelke, E. (2006). Core knowledge of geometry in an Amazonian indigene group. *Science*, 311(5759):381–384.

Delahaye, J.-P. and Zenil, H. (2012). Numerical evaluation of algorithmic complexity for short strings: A glance into the innermost structure of randomness. *Applied Mathematics and Computation*, 219(1):63–77.

Delaherche, E., Chetouani, M., Mahdhaoui, A., Saint-Georges, C., Viaux, S., and Cohen, D. (2012). Interpersonal synchrony: A survey of evaluation methods across disciplines. *IEEE Transactions on Affective Computing*, 3(3):349–365.

Demasi, P., Szwarcfiter, J. L., and Cruz, A. J. O. (2010). A theoretical framework to formalize AGI-Hard problems. In Hutter, M., Baum, E., and Kitzelmann, E. (Eds.), *3rd Conference on Artificial General Intelligence*, pages 178–179. Atlantis Press.

Dennett, D. (1988). Conditions of personhood. In Goodman, M. F., editor, *What is a person? Contemporary issues in biomedicine, ethics, and society*, pages 145–167. Humana Press.

Dennett, D. C. (1989). *The intentional stance*. MIT Press.

Dennett, D. C. (1995). Darwin's dangerous idea. *The Sciences*, 35(3):34–40.

Dennett, D. C. (1996). *Kinds of minds: Toward an understanding of consciousness.* Basic Books.

Detterman, D. K. (1982). Does "g" exist? *Intelligence*, 6(2):99–108.

Detterman, D. K. (1991). Reply to Deary and Pagliari: Is g intelligence or stupidity? *Intelligence*, 15(2):251–255.

Detterman, D. K. (2011). A challenge to Watson. *Intelligence*, 39(2-3):77–78.

Detterman, D. K. and Daniel, M. H. (1989). Correlations of mental tests with each other and with cognitive variables are highest for low IQ groups. *Intelligence*, 13(4):349–359.

Detterman, D. K., Gabriel, L. T., and Ruthsatz, J. M. (2000). Intelligence and mental retardation. In Sternberg, R., editor, *Handbook of intelligence*, pages 141–158. Cambridge University Press.

Deutsch, D. (1985). Quantum theory, the Church-Turing principle and the universal quantum computer. *Proceedings of the Royal Society of London, Series A*, 400(1818):97–117.

Deutsch, D. (2012). How close are we to creating artificial intelligence? https://aeon.co/essays/how-close-are-we-to-creating-artificial-intelligence.

Dewey, D. (2011). Learning what to value. In J. Schmidhuber, K. R. Thórisson, and M. Looks (Eds.), *4th International Conference on Artificial General Intelligence*, pages 309–314. Springer.

Dickerson, J. P., Kagan, V., and Subrahmanian, V. (2014). Using sentiment to detect bots on Twitter: Are humans more opinionated than bots? In *Advances in Social Networks Analysis and Mining (ASONAM)*, pages 620–627. IEEE.

Didier, C., Duan, W., Dupuy, J.-P., Guston, D. H., Liu, Y., López, C. J., Michelfelder, D., Mitcham, C., Sarewitz, D., Stilgoe, J., et al. (2015). Acknowledging AI's dark side. *Science*, 349(6252):1064.

Dimitrakakis, C., Li, G., and Tziortziotis, N. (2014). The reinforcement learning competition 2014. *AI Magazine*, 35(3):61–65.

Dimou, C., Symeonidis, A. L., Mitkas, P., et al. (2007). Towards a generic methodology for evaluating MAS performance. In *International Conference on Integration of Knowledge Intensive Multi-Agent Systems, 2007 (KIMAS 2007)*, pages 174–179. IEEE.

Dinnerstein, D. and Wertheimer, M. (1957). Some determinants of phenomenal overlapping. *American Journal of Psychology*, 70:21–37.

Dowe, D. L. (2008). Foreword re C. S. Wallace. *Computer Journal*, 51(5):523–560. Christopher Stewart WALLACE (1933-2004) memorial special issue.

Dowe, D. L. (2011). MML, hybrid Bayesian network graphical models, statistical consistency, invariance and uniqueness. In P. S. Bandyopadhyay and M. R. Forster, editors, *Handbook of the philosophy of science, Volume 7: Philosophy of statistics*, pages 901–982. Elsevier.

Dowe, D. L. (2013). Introduction to Ray Solomonoff 85th memorial conference. In Dowe, D. L., editor, *Algorithmic probability and friends: Bayesian prediction and artificial intelligence*, pages 1–36. *Lecture Notes in Computer Science*, volume 7070. Springer.

Dowe, D. L. and Hajek, A. R. (1997). A computational extension to the Turing test. In *Proceedings of the 4th Conference of the Australasian Cognitive Science Society, University of Newcastle, NSW, Australia*.

Dowe, D. L. and Hajek, A. R. (1998). A non-behavioural, computational extension to the Turing Test. In *International Conference on Computational Intelligence and Multimedia Applications (ICCIMA'98), Gippsland, Australia*, pages 101– 106.

Dowe, D. L. and Hernández-Orallo, J. (2012). IQ tests are not for machines, yet. *Intelligence*, 40(2):77–81.

Dowe, D. L. and Hernández-Orallo, J. (2014). How universal can an intelligence test be? *Adaptive Behavior*, 22(1):51–69.

Dowe, D. L., Hernández-Orallo, J., and Das, P. K. (2011). Compression and intelligence: Social environments and communication. In Schmidhuber, J., Thórisson, K., and Looks, M., editors, 4th International Conference on *Artificial general intelligence*, pages 204–211. *Lecture Notes in Artificial Intelligence*, volume 6830. Springer.

Downing, S. M. and Haladyna, T. M. (2006). *Handbook of test development.* Lawrence Erlbaum Associates.

Drescher, G. L. (1991). *Made-up minds: A constructivist approach to artificial intelligence*. MIT Press.

Drexler, E. K. (1986). *Engines of creation: The coming era of nanotechnology*. Anchor Books.

Drummond, C. and Japkowicz, N. (2010). Warning: Statistical benchmarking is addictive. Kicking the habit in machine learning. *Journal of Experimental and Theoretical Artificial Intelligence*, 22(1):67–80.

Dunér, D. (2011). Astrocognition: Prolegomena to a future cognitive history of exploration. In *Humans in outer space: Interdisciplinary Perspectives*, pages 117–140. Springer.

Dunn, L. M. and Dunn, D. (2012). *Peabody Picture Vocabulary Test, (PPVT-4)*. Pearson.

Durfee, E. H. and Montgomery, T. A. (1990). MICE: A flexible testbed for intelligent coordination experiments. In Erman, L., editor, *Intelligent real time problem solving: Workshop report*, pages X1–X16.

Dusenbery, D. B. (1992). *Sensory ecology*. Freeman.

Dutton, E. and Lynn, R. (2013). A negative Flynn effect in Finland, 1997–2009. *Intelligence*, 41(6):817–820.

Eaton, E., Lane, T., et al. (2008). Modeling transfer relationships between learning tasks for improved inductive transfer. In *Machine learning and knowledge discovery in databases*, pages 317–332. Springer.

Eden, A. H. (2016). The singularity controversy. Technical Report, STR 2016-1. doi:10.13140/RG.2.1.3416.6809.

Eden, A. H., Moor, J. H., Soraker, J. H., and Steinhart, E. (2012a). *Singularity hypotheses: A scientific and philosophical assessment*. Springer.

Eden, A. H., Steinhart, E., Pearce, D., and Moor, J. H. (2012b). Singularity hypotheses: An overview. In Eden, A. H., Moor, J. H., Soraker, J. H., and Steinhart, E., editors, *Singularity hypotheses*, pages 1–12. Springer.

Edmonds, B. (2009). The social embedding of intelligence. In Epstein, R., Roberts, G., and Beber, G., editors, *Parsing the Turing test*, pages 211–235. Springer.

Edmondson, W. (2012). The intelligence in ETI – What can we know? *Acta Astronautica*, 78:37–42.

Eliasmith, C., Stewart, T. C., Choo, X., Bekolay, T., DeWolf, T., Tang, C., and Rasmussen, D. (2012). A large-scale model of the functioning brain. *Science*, 338(6111):1202–1205.

Elo, A. E. (1978). *The rating of chessplayers, past and present*, volume 3. Batsford.

Embretson, S. E. (2004). The second century of ability testing: Some predictions and speculations. *Measurement: Interdisciplinary Research and Perspectives*, 2(1):1–32.

Embretson, S. E. and McCollam, K. M. S. (2000). Psychometric approaches to understanding and measuring intelligence. In Sternberg, R., editor, *Handbook of intelligence*, pages 423–444. Cambridge University Press.

Embretson, S. E. and Reise, S. P. (2000). *Item response theory for psychologists*. Lawrence Erlbaum Associates.

Emery, N. J., Clayton, N. S., and Frith, C. D. (2007). Introduction. Social intelligence: from brain to culture. *Philosophical Transactions of the Royal Society, Series B*, 362(1480):485–488.

Engel, D., Woolley, A. W., Aggarwal, I., Chabris, C. F., Takahashi, M., Nemoto, K., Kaiser, C., Kim, Y. J., and Malone, T. W. (2015). Collective intelligence in computer-mediated collaboration emerges in different contexts and cultures. In *Conference on Human Factors in Computing Systems*, pages 3769–3778. ACM.

Engelbart, D. C. (1962). Augmenting human intellect: A conceptual framework. Summary Report, AFOSR-3233. Stanford Research Institute.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., and Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, 128(3):309.

Epstein, R., Roberts, G., and Beber, G., editors (2008). *Parsing the Turing test: Philosophical and methodological issues in the quest for the thinking computer*. Springer.

Erickson, G. J. and Smith, C. R. (1988). *Maximum-entropy and Bayesian methods in science and engineering*, volume 31. Springer Science and Business Media.

Ernst, G. W. and Newell, A. (1969). *GPS: A case study in generality and problem solving*. Academic Press.

EU (2010). Directive 2010/63/EU of the European Parliament and of the Council of 22 september 2010 on the protection of animals used for scientific purposes. *Official Journal of the European Union*, 276:33–74.

Evans, J. M. and Messina, E. R. (2001). Performance metrics for intelligent systems. *NIST Special Publication SP*, pages 101–104.

Evans, N. and Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, 32(05):429–448.

Evans, T. G. (1963). *A heuristic program of solving geometric analogy problems*. PhD thesis, MIT.

Evans, T. G. (1965). A heuristic program to solve geometric-analogy problems. In *Proceedings Spring Joint Computer Conference*, volume 25, pages 327–339. Spartan Books.

Everitt, T., Lattimore, T., and Hutter, M. (2014). Free lunch for optimisation under the universal distribution. In *2014 IEEE Congress on Evolutionary Computation (CEC)*, pages 167–174. IEEE.

Eysenck, H. J. (2007). *The structure and measurement of intelligence*. Transaction.

Eysenck, H. J., Eysenck, M. W., Ellis, A., Hunt, E., and Johnson-Laird, P. (1985). *Personality and individual differences: A natural science approach*. Plenum.

Eysenck, H. J. and Eysenck, S. B. G. (1975). *Manual of the Eysenck Personality Questionnaire (junior and adult)*. Hodder and Stoughton.

Falkenauer, E. (1998). On method overfitting. *Journal of Heuristics*, 4(3):281–287.

Falmagne, J.-C. and Doignon, J.-P. (2010). *Learning spaces: Interdisciplinary applied mathematics*. Springer Science and Business Media.

Farah, M. J. and Heberlein, A. S. (2007). Personhood and neuroscience: Naturalizing or nihilating? *American Journal of Bioethics*, 7(1):37–48.

Federle, M. J., Bassler, B. L., et al. (2003). Interspecies communication in bacteria. *Journal of Clinical Investigation*, 112(9):1291–1299.

Feigenbaum, E. A., Buchanan, B. G., and Lederberg, J. (1970). *On generality and problem solving: A case study using the DENDRAL program*. Stanford University.

Feil-Seifer, D., Skinner, K., and Matarić, M. J. (2007). Benchmarks for evaluating socially assistive robotics. *Interaction Studies*, 8(3):423–439.

Feldman, J. (2000). Minimization of boolean complexity in human concept learning. *Nature*, 407(6804):630–633.

Feldman, J. (2003). Simplicity and complexity in human concept learning. *The General Psychologist*, 38(1):9–15.

Feng, A. and Graetz, G. (2015). *Rise of the machines: The effects of labor-saving innovations on jobs and wages*. Centre for Economic Performance, London School of Economics and Political Science.

Fenson, L. (2007). *MacArthur-Bates Communicative Development Inventories: User's guide and technical manual*. Paul H. Brookes.

Fernandes, H. B., Woodley, M. A., and te Nijenhuis, J. (2014). Differences in cognitive abilities among primates are concentrated on g: Phenotypic and phylogenetic comparisons with two meta-analytical databases. *Intelligence*, 46:311–322.

Ferns, N., Panangaden, P., and Precup, D. (2004). Metrics for finite Markov decision processes. In *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence*, pages 162–169. AUAI Press.

Ferrando, P. J. (2009). Difficulty, discrimination, and information indices in the linear factor analysis model for continuous item responses. *Applied Psychological Measurement*, 33(1):9–24.

Ferrando, P. J. (2012). Assessing the discriminating power of item and test scores in the linear factor-analysis model. *Psicológica*, 33:111–139.

Ferrara, E., Varol, O., Davis, C., Menczer, F., and Flammini, A. (2015). The rise of social bots. CoRR arXiv preprint:1407.5225.

Ferrucci, D., Brown, E., Chu-Carroll, J., Fan, J., Gondek, D., Kalyanpur, A. A., Lally, A., Murdock, J., Nyberg, E., Prager, J., et al. (2010). Building Watson: An overview of the DeepQA project. *AI Magazine*, 31(3):59–79.

Fidora, A. and Sierra, C. E. (2011). *Ramon Llull: From the Ars Magna to artificial intelligence*. Artificial Intelligence Research Institute, IIIA.

Fiore, V. G., Mannella, F., Mirolli, M., Gurney, K., and Baldassarre, G. (2008). Instrumental conditioning driven by apparently neutral stimuli: A model tested with a simulated robotic rat. In *International Conference on Epigenetic Robotics: Modeling cognitive development in robotic systems*, pages 13–20. IEEE.

Fiorito, G. and Scotto, P. (1992). Observational learning in octopus vulgaris. *Science*, 256(5056):545–547.

Firn, R. (2004). Plant intelligence: An alternative point of view. *Annals of Botany*, 93(4):345–351.

Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, 87(6):477.

Flach, P. (2012). *Machine learning: The art and science of algorithms that make sense of data*. Cambridge University Press.

Fletcher, J. F. (1979). *Humanhood: Essays in biomedical ethics*. Prometheus Books.

Fleuret, F., Li, T., Dubout, C., Wampler, E. K., Yantis, S., and Geman, D. (2011). Comparing machines and humans on a visual categorization test. *Proceedings of the National Academy of Sciences*, 108(43):17621–17625.

Floridi, L. (2013). *The ethics of information*. Oxford University Press.

Floridi, L. and Sanders, J. W. (2004). On the morality of artificial agents. *Minds and Machines*, 14(3):349–379.

Flynn, J. R. (2007). *What is intelligence? Beyond the Flynn effect*. Cambridge University Press.

Flynn, J. R. (2013). The “Flynn effect” and Flynn’s paradox. *Intelligence*, 41(6):851–857.

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. MIT Press.

Fogarty, G. J. and Stankov, L. (1995). Challenging the “law of diminishing returns”. *Intelligence*, 21(2):157–174.

Fogel, D. B. (1995). *Evolutionary computation: Toward a new philosophy of machine intelligence*. IEEE.

Ford, K. M. and Hayes, P. J. (1998). On computational wings: Rethinking the goals of artificial intelligence – the gold standard of traditional artificial intelligence – passing the so-called Turing test and thereby appearing to be. *Scientific American*, 9(Winter):78–83.

Fostel, G. (1993). The Turing test is for the birds. *ACM SIGART Bulletin*, 4(1): 7–8.

Fountain, S. B. (2006). The structure of sequential behavior. In Wasserman, E. A. and Zentall, T. R., editors, *Comparative cognition: Experimental explorations of animal intelligence*, pages 439–458. Oxford University Press.

Fowler, O. S. and Fowler, L. N. (1859). *Self-instructor in phrenology and physiology*. Fowler and Wells.

Fragaszy, D., Johnson-Pynn, J., Hirsh, E., and Brakke, K. (2003). Strategic navigation of two-dimensional alley mazes: Comparing capuchin monkeys and chimpanzees. *Animal Cognition*, 6:149–160.

Fragaszy, D. M., Kennedy, E., Murnane, A., Menzel, C., Brewer, G., Johnson-Pynn, J., and Hopkins, W. (2009). Navigating two-dimensional mazes: Chimpanzees (*Pan troglodytes*) and capuchins (*Cebus apella* sp.) profit from experience differently. *Animal Cognition*, 12(3):491–504.

Frankfurt, H. G. (1971). Freedom of the will and the concept of a person. *Journal of Philosophy*, 68(1):5–20.

French, R. M. (2000). The Turing test: The first 50 years. *Trends in Cognitive Sciences*, 4(3):115–122.

French, R. M. (2012). Moving beyond the Turing test. *Communications of the ACM*, 55(12):74–77.

Frey, C. B. and Osborne, M. A. (2013). *The future of employment: How susceptible are jobs to computerisation?* University of Oxford.

Frith, U. and Frith, C. (2010). The social brain: Allowing humans to boldly go where no other species has been. *Philosophical Transactions of the Royal Society of London*, Series B, 365(1537):165–176.

Gabrieli, J. D., Ghosh, S. S., and Whitfield-Gabrieli, S. (2015). Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience. *Neuron*, 85(1):11–26.

Gács, P. and Vitányi, P. M. (2011). Raymond J. Solomonoff 1926–2009. *IEEE Information Theory Society Newsletter*, 61(1):11–16.

Galperin, M. Y. (2005). A census of membrane-bound and intracellular signal transduction proteins in bacteria: Bacterial IQ, extroverts and introverts. *BMC Microbiology*, 5(1):35.

Galton, F. (1869). *Hereditary genius: An Inquiry into its laws and consequences*, 3rd edition 1892. Macmillan.

Galton, F. (1883). *Inquiries into the human faculty and its development*. Macmillan.

Galton, F. (1884). Measurement of character. *Fortnightly Review*, 36:179–185.

Galton, F. (1906). Cutting a round cake on scientific principles (letters to the editor). *Nature*, 75(1938):173.

Galton, F. (1907). Vox populi. *Nature*, 75:450–451.

García, K. (2013). Asume que no es humano al ser incapaz de reconocer un código CAPTCHA. *El Mundo Today*. http://www.elmundotoday.com/2013/03/asume-que-no-es-humano-al-ser-incapaz-de-reconocer-un-codigo-captcha/.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. Basic Books.

Gardner, R. A. and Gardner, B. T. (1969). Teaching sign language to a chimpanzee. *Science*, 165(3894):664–672.

Gauvrit, N., Zenil, H., and Tegnér, J. (2015). The information-theoretic and algorithmic approach to human, animal and artificial cognition. CoRR arXiv preprint:1501.04242.

Gavane, V. (2013). A measure of real-time intelligence. *Journal of Artificial General Intelligence*, 4(1):31–48.

Gelman, S. A., & Kalish, C. W. (2006). Conceptual development. In D. Kuhn & R. Siegler (Eds.), *Handbook of child psychology*, Vol. 2: *Cognition, perception and language*. pages 687–733. John Wiley.

Genesereth, M., Love, N., and Pell, B. (2005). General game playing: Overview of the AAAI competition. *AI Magazine*, 26(2):62.

Genesereth, M. and Thielscher, M. (2014). General game playing. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 8(2):1–229.

Gentner, T. Q., Fenn, K. M., Margoliash, D., and Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature*, 440(7088):1204–1207.

Gerbino, W. and Salmaso, D. (1987). The effect of amodal completion on visual matching. *Acta Psychologica*, 65(1):25–46.

Gerónimo, D. and López, A. M. (2014). Datasets and benchmarking. In *Vision-based pedestrian protection systems for intelligent vehicles*, pages 87–93. Springer.

Gibson, W. (1984). *Neuromancer*. Penguin.

Ginsberg, M. (1993). *Essentials of artificial intelligence*. Newnes.

Ginsburg, S. and Jablonka, E. (2009). Epigenetic learning in non-neural organisms. *Journal of Biosciences*, 34(4):633–646.

Glaser, R. (1963). Instructional technology and the measurement of learing outcomes: Some questions. *American Psychologist*, 18(8):519.

Glasersfeld, E. v. (1974). The Yerkish language for non-human primates. *American Journal of Computational Linguistics*, 1.

Goertzel, B. (2006). *The hidden pattern*. Brown Walker.

Goertzel, B. (2010). Toward a formal characterization of real-world general intelligence. In Hutter, M., Baum, E., and Kitzelmann, E., editors, 3rd International Conference on Artificial General Intelligence, pages 19–24. Atlantis Press.

Goertzel, B., Arel, I., and Pennachin, C. (2010). Cogbot: An integrative cognitive architecture aimed at emulating early childhood intelligence in a humanoid robot. http://www.goertzel.org/CogBot_public.pdf.

Goertzel, B., Arel, I., and Scheutz, M. (2009). *Toward a roadmap for human-level artificial general intelligence: Embedding HLAI systems in broad, approachable, physical or virtual contexts*. Artificial General Intelligence Roadmap Initiative.

Goertzel, B. and Bugaj, S. V. (2009). AGI preschool: A framework for evaluating early-stage human-like AGIs. In *the 2nd International Conference on Artificial General Intelligence (AGI-09)*, pages 31–36. Atlantis Press.

Goertzel, B. and Pennachin, C., editors (2007). *Artificial general intelligence*. Springer.

Golbeck, J., Robles, C., Edmondson, M., and Turner, K. (2011). Predicting personality from Twitter. In *Privacy, security, risk and trust: 2011 IEEE Third International Conference on Social Computing (socialcom)*, pages 149–156. IEEE.

Gold, E. M. (1967). Language identification in the limit. *Information and Control*, 10(5):447–474.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., and Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40(1):84–96.

Goldin, D., Smolka, S. A., and Wegner, P. (2006). *Interactive computation*. Springer.

Goldreich, O. and Vadhan, S. (2007). Special issue on worst-case versus average-case complexity – editors' foreword. *Computational Complexity*, 16(4):325–330.

Goleman, D. (1996). Emotional intelligence: Why it can matter more than IQ. *Learning*, 24(6):49–50.

Goleman, D. (2007). *Social intelligence*. Random House.

Good, I. J. (1965). Speculations concerning the first ultraintelligent machine. *Advances in Computers*, 6(99):31–83.

Good, I. J. (1971). Twenty-seven principles of rationality. In *Foundations of statistical inference*, page 197. Holt, Rinehart, Winston.

Gordon, B. B. (2007). Report on panel discussion on (re-)establishing or increasing collaborative links between artificial intelligence and intelligent systems. In Messina, E. R. and Madhavan, R., editors, *Proceedings of the 2007 Workshop on Performance Metrics for Intelligent Systems*, pages 302–303. NIST.

Gosling, S. D. (2001). From mice to men: What can we learn about personality from animal research? *Psychological Bulletin*, 127(1):45.

Gosling, S. D. and John, O. P. (1999). Personality dimensions in nonhuman animals a cross-species review. *Current Directions in Psychological Science*, 8(3):69–75.

Gottfredson, L. S. (1997a). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, 24(1):13–23.

Gottfredson, L. S. (1997b). Why g matters: The complexity of everyday life. *Intelligence*, 24(1):79–132.

Gould, S. J. (1981). *The mismeasure of man*, revised and expanded edition. W. W. Norton.

Grassé, P. P. (1959). La reconstruction du nid et les coordinations interindividuelles chez bellicositermes natalensis et cubitermes sp. la théorie de la stigmergie: Essai d'interprétation du comportement des termites constructeurs. *Insectes sociaux*, 6(1):41–80.

Greely, H., Sahakian, B., Harris, J., Kessler, R. C., Gazzaniga, M., Campbell, P., and Farah, M. J. (2008). Towards responsible use of cognitive-enhancing drugs by the healthy. *Nature*, 456(7223):702–705.

Griffiths, R. (1954). *The abilities of babies: A study in mental measurement.* McGraw-Hill.

Grigorenko, E. L. and Sternberg, R. J. (1998). Dynamic testing. *Psychological Bulletin*, 124(1):75.

Groeger, J. A., Field, D., and Hammond, S. M. (1999). Measuring memory span. *International Journal of Psychology*, 34(5–6):359–363.

Grotzer, T. A. and Perkins, D. (2004). Teaching intelligence: A performance conception. In Sternberg, R., editor, *Handbook of intelligence*, pages 492–518. Cambridge University Press.

Grünwald, P. D., Myung, I. J., and Pitt, M. A., editors (2005). *Advances in minimum description length: Theory and applications*. MIT Press.

Guilford, J. (1982). Cognitive psychology's ambiguities: Some suggested remedies. *Psychological Review*, 89(1):48.

Guilford, J. P. (1956). The structure of intellect. *Psychological Bulletin*, 53(4):267.

Guilford, J. P. (1967). *The nature of human intelligence*. McGraw-Hill.

Gulliksen, H. (1945). The relation of item difficulty and inter-item correlation to test variance and reliability. *Psychometrika*, 10(2):79–91.

Gulwani, S., Hernández-Orallo, J., Kitzelmann, E., Muggleton, S. H., Schmid, U., and Zorn, B. (2015). Inductive programming meets the real world. *Communications of the ACM*, 58(11):90–99.

Gunkel, D. J. (2012). *The machine question: Critical perspectives on AI, robots, and ethics*. MIT Press.

Gunkel, D. J. (2014). A vindication of the rights of machines. *Philosophy and Technology*, 27(1):113–132.

Guttman, L. (1954). *A new approach to factor analysis: The Radex*. Free Press.

Hahn, U., Chater, N., and Richardson, L. B. (2003). Similarity as transformation. *Cognition*, 87(1):1–32.

Hall, J. S. (2007). *Beyond AI: Creating the conscience of the machine*. Prometheus Books.

Halmes, M. (2013). Measurements of collective machine intelligence. CoRR arXiv preprint:1306.6649.

Halpin, H., Clark, A., and Wheeler, M. (2010). *Towards a philosophy of the web: Representation, enaction, collective intelligence*. Web Science Conference.

Hamilton, W. (1859). *Lectures on metaphysics and logic*, volume 1. Gould and Lincoln.

Hammill, D. D. and Newcomer, P. L. (2008). *Test of language development – intermediate (TOLD I: 3)*. Pro-Ed.

Hampson, R. E., Song, D., Chan, R. H., Sweatt, A. J., Riley, M. R., Gerhardt, G., Shin, D. C., Marmarelis, V. Z., Berger, T. W., Deadwyler, S., et al. (2012). A nonlinear model for hippocampal cognitive prosthesis: Memory facilitation by hippocampal ensemble stimulation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 20(2):184–197.

Hampson, R. E., Song, D., Opris, I., Santos, L. M., Shin, D. C., Gerhardt, G. A., Marmarelis, V. Z., Berger, T. W., and Deadwyler, S. A. (2013). Facilitation of memory encoding in primate hippocampus by a neuroprosthesis that promotes task-specific neural firing. *Journal of Neural Engineering*, 10(6):066013.

Hand, D. J. (2004). *Measurement theory and practice*. Hodder Arnold.

Haney, R. (1969). Classical conditioning of a plant: Mimosa pudica. *Journal of Biological Psychology*, 11:5–12.

Hanks, S., Pollack, M. E., and Cohen, P. R. (1993). Benchmarks, test beds, controlled experimentation, and the design of agent architectures. *AI Magazine*, 14(4): 17.

Hanus, D., Mendes, N., Tennie, C., and Call, J. (2011). Comparing the performances of apes (*Gorilla gorilla, Pan troglodytes, Pongo pygmaeus*) and human children (*Homo sapiens*) in the floating peanut task. *PloS ONE*, 6(6):e19555.

Haraway, D. J. (1991). *Simians, cyborgs, and women: The reinvention of nature*. Routledge.

Hardman, D. K. and Payne, S. J. (1995). Problem difficulty and response format in syllogistic reasoning. *Quarterly Journal of Experimental Psychology*, 48(4):945–975.

Hare, B. and Tomasello, M. (2005). Human-like social skills in dogs? *Trends in Cognitive Sciences*, 9(9):439–444.

Harnad, S. (1992). The Turing Test is not a trick: Turing indistinguishability is a scientific criterion. *ACM SIGART Bulletin*, 3(4):9–10.

Harper, J. L., et al. (1977). *Population biology of plants.* Academic Press.

Hartley, R. V. L. (1928). Transmission of information. *Bell System Technical Journal*, 7(3):535–563.

Hartmann, P. and Teasdale, T. W. (2004). A test of Spearman's "law of diminishing returns" in two large samples of Danish military draftees. *Intelligence*, 32(5):499–508.

Hartshorne, J. K. and Germine, L. T. (2015). When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. *Psychological Science*, 26(4):433–443.

Hauser, M. D., Barner, D., and O'Donnell, T. (2007). Evolutionary linguistics: A new look at an old landscape. *Language Learning and Development*, 3(2):101–132.

Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298(5598):1569–1579.

Havel, I. M. (2013). On the way to intelligence singularity. In *Beyond artificial intelligence*, pages 3–26. Springer.

Hayes, B. (2002). Computing science: The easiest hard problem. *American Scientist*, pages 113–117.

Hayes, P. and Ford, K. (1995). Turing test considered harmful. In *International Joint Conference on Artificial Intelligence (IJCAI)*, pages 972–977. IJCAI.

Haywood, H. C. (1988). Dynamic assessment: The learning potential assessment device. In R. Jones (Ed.), *Psychoeducational assessment of minority group children: A casebook*, pages 39–63. Cobb and Henry.

He, J., Reeves, C., Witt, C., and Yao, X. (2007). A note on problem difficulty measures in black-box optimization: Classification, realizations and predictability. *Evolutionary Computation*, 15(4):435–443.

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

Herculano-Houzel, S., Catania, K., Manger, P. R., and Kaas, J. H. (2015). Mammalian brains are made of these: A dataset of the numbers and densities of neuronal and nonneuronal cells in the brain of glires, primates, scandentia, eulipotyphlans, afrotherians and artiodactyls, and their relationship with body mass. *Brain, Behavior, and Evolution*, 86(3–4):145–163.

Hermann, K. M., Kočiskỳ, T., Grefenstette, E., Espeholt, L., Kay, W., Suleyman, M., and Blunsom, P. (2015). Teaching machines to read and comprehend. CoRR arXiv preprint:1506.03340.

Hermans, D., Knoors, H., and Verhoeven, L. (2009). Assessment of sign language development: The case of deaf children in the Netherlands. *Journal of Deaf Studies and Deaf Education*, 15(2):107–119.

Hernández-Orallo, J. (1999a). *Computational measures of information gain and reinforcement in inference processes*. PhD thesis, Department of Logic and Philosophy of Science, University of Valencia. http://users.dsic.upv.es/~jorallo/tesi/.

Hernández-Orallo, J. (1999b). Universal and cognitive notions of part. In *Ferrer, L., et al., editor, Proceedings of 4th Systems Science European Congress*, pages 711–722. Sociedad Española de Sistemas Generales.

Hernández-Orallo, J. (2000a). Beyond the Turing test. *Journal of Logic, Language, and Information*, 9(4):447–466.

Hernández-Orallo, J. (2000b). A computational definition of 'consilience'. *Philosophica*, 61:901–920.

Hernández-Orallo, J. (2000c). Constructive reinforcement learning. *International Journal of Intelligent Systems*, 15(3):241–264.

Hernández-Orallo, J. (2000d). On the computational measurement of intelligence factors. In Meystel, A. M. and Messina, E. R., editors, *Measuring the performance and intelligence of systems: Proceedings of the 2000 PerMIS Workshop*, pages 72–79. NIST Special Publication 970. NIST.

Hernández-Orallo, J. (2000e). Thesis: Computational measures of information gain and reinforcement in inference processes. *AI Communications*, 13(1):49–50.

Hernández-Orallo, J. (2010). A (hopefully) non-biased universal environment class for measuring intelligence of biological and artificial systems. In Hutter, M., Baum, E.,

and Kitzelmann, E., editors, *Third International Conference on Artificial General Intelligence*, pages 182–183. Atlantis Press.

Hernández-Orallo, J. (2015a). C-tests revisited: Back and forth with complexity. In Bieger, J., Goertzel, B., and Potapov, A., editors, *Artificial General Intelligence – 8th International Conference, AGI 2015, Berlin, Germany, July 22–25, 2015, Proceedings*, pages 272–282. Springer.

Hernández-Orallo, J. (2015b). A note about the generalisation of the C-tests. arXiv preprint:1412.8529.

Hernández-Orallo, J. (2015c). On environment difficulty and discriminating power. *Autonomous Agents and Multi-Agent Systems*, 29:402–454.

Hernández-Orallo, J. (2015d). Stochastic tasks: Difficulty and Levin search. In Bieger, J., Goertzel, B., and Potapov, A., editors, *Artificial General Intelligence – 8th International Conference, AGI 2015, Berlin, Germany, July 22–25, 2015, Proceedings*, pages 90–100. Springer.

Hernández-Orallo, J. (2015e). Universal psychometrics tasks: Difficulty, composition and decomposition. arXiv preprint:1503.07587.

Hernández-Orallo, J. (2016). Evaluation in artificial intelligence: From task-oriented to ability-oriented measurement. *Artificial Intelligence Review*. doi: http://dx.doi.org/10.1007/s10462-016-9505-7.

Hernández-Orallo, J. and Dowe, D. L. (2010). Measuring universal intelligence: Towards an anytime intelligence test. *Artificial Intelligence*, 174(18):1508–1539.

Hernández-Orallo, J. and Dowe, D. L. (2013). On potential cognitive abilities in the machine kingdom. *Minds and Machines*, 23:179–210.

Hernández-Orallo, J., Dowe, D. L., España-Cubillo, S., Hernández-Lloreda, M. V., and Insa-Cabrera, J. (2011). On more realistic environment distributions for defining, evaluating and developing intelligence. In Schmidhuber, J., Thórisson, K., and Looks, M., editors, 4th International Conference on *A*rtificial general intelligence, pages 82–91. *Lecture Notes in Artificial Intelligence*, volume 6830. Springer.

Hernández-Orallo, J., Dowe, D. L., and Hernández-Lloreda, M. V. (2014). Universal psychometrics: Measuring cognitive abilities in the machine kingdom. *Cognitive Systems Research*, 27:50–74.

Hernández-Orallo, J. and García-Varea, I. (2000). Explanatory and creative alternatives to the MDL principle. *Foundations of Science*, 5(2):185–207.

Hernández-Orallo, J., Insa-Cabrera, J., Dowe, D. L., and Hibbard, B. (2012). Turing tests with Turing machines. In Voronkov, A., editor, *Turing-100*, volume 10, pages 140–156. EPiC Series.

Hernández-Orallo, J., Martínez-Plumed, F., Schmid, U., Siebers, M., and Dowe, D. L. (2016). Computer models solving human intelligence test problems: Progress and implications. *Artificial Intelligence*, 230:74–107.

Hernández-Orallo, J. and Minaya-Collado, N. (1998). A formal definition of intelligence based on an intensional variant of Kolmogorov complexity. In *Proceedings of the International Symposium of Engineering of Intelligent Systems (EIS'98)*, pages 146–163. ICSC Press.

Herrmann, E. and Call, J. (2012). Are there geniuses among the apes? *Philosophical Transactions of the Royal Society, Series B*, 367(1603):2753–2761.

Herrmann, E., Call, J., Hernández-Lloreda, M. V., Hare, B., and Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science*, 317(5843):1360–1366.

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

Herrmann, E., Misch, A., Hernandez-Lloreda, V., and Tomasello, M. (2015). Uniquely human self-control begins at school age. *Developmental Science*, 18(6):979–993.

Herrnstein, R. J. and Murray, C. (1994). *Bell curve: Intelligence and class structure in American life*. Simon and Schuster.

Herzing, D. L. (2014). Profiling nonhuman intelligence: An exercise in developing unbiased tools for describing other "types" of intelligence on Earth. *Acta Astronautica*, 94(2):676–680.

Heylighen, F. (2011). Conceptions of a global brain: An historical review. In Grinin, L. E., Carneiro, R. L., Korotayev A. V., and Spier F., editors, *Evolution: Cosmic, biological, and social*, pages 274–289. Uchitel.

Hibbard, B. (2002). Humans will want to become super-intelligent machines. In *Super-intelligent machines*, pages 193–198. Springer.

Hibbard, B. (2008). Adversarial sequence prediction. *Frontiers in Artificial Intelligence and Applications*, 171:399.

Hibbard, B. (2009). Bias and no free lunch in formal measures of intelligence. *Journal of Artificial General Intelligence*, 1(1):54–61.

Hibbard, B. (2011). Measuring agent intelligence via hierarchies of environments. In Schmidhuber, J., Thórisson, K., and Looks, M., editors, 4th International Conference on *A*rtificial general intelligence, pages 303–308. *Lecture Notes in Artificial Intelligence*, volume 6830. Springer.

Hibbard, B. (2015a). *Bill Hibbard on ethical artificial intelligence, an interview by Luke Muehlhauser*. Machine Intelligence Research Institute.

Hibbard, B. (2015b). Ethical artificial intelligence. CoRR arXiv preprint:1411.1373.

Hingston, P. (2010). A new design for a Turing test for bots. In *2010 IEEE Symposium on Computational Intelligence and Games (CIG)*, pages 345–350. IEEE.

Hingston, P. (2012). *Believable bots: Can computers play like people?* Springer.

Hirschberg, J. and Manning, C. D. (2015). Advances in natural language processing. *Science*, 349(6245):261–266.

Ho, T. K. and Basu, M. (2002). Complexity measures of supervised classification problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(3):289–300.

Hoen, P. J., Tuyls, K., Panait, L., Luke, S., and La Poutre, J. A. (2005). An overview of cooperative and competitive multiagent learning. In *International Conference on Learning and Adaption in Multi-Agent Systems*, pages 1–46. Springer.

Hoff, E. (2013). *Language development*. Cengage Learning.

Hofstadter, D. R. (1979). *Gödel, Escher, Bach: An eternal golden braid*. Basic Books.

Holekamp, K. E. (2007). Questioning the social intelligence hypothesis. *Trends in Cognitive Sciences*, 11(2):65–69.

Holtgraves, T. M. (2014). *The Oxford handbook of language and social psychology*. Oxford University Press.

Holz, D., Iocchi, L., and Van Der Zant, T. (2013). Benchmarking intelligent service robots through scientific competitions: The RoboCup@Home approach. In *AAAI Spring Symposium: Designing Intelligent Robots*, pp. 27–32. Association for the Advancement of Artificial Intelligence.

Hong, L. and Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences of the United States of America*, 101(46):16385–16389.

Hoos, H. H. (1999). SAT-encodings, search space structure, and local search performance. In *International Joint Conference on Artificial Intelligence (IJCAI)*, volume 99, pages 296–303. IJCAI.

Hopcroft, J. E., Motwani, R., and Ullman, J. D. (2007). *Introduction to automata theory, languages, and computation*, 3rd edition. Pearson.

Horn, J. L. and Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica*, 26:107–129.

Hortal, J., Araújo, M. B., and Lobo, J. M. (2009). Testing the effectiveness of discrete and continuous environmental diversity as a surrogate for species diversity. *Ecological Indicators*, 9(1):138–149.

Howe, J. (2008). *Crowdsourcing: How the power of the crowd is driving the future of business*. Random House.

Huarte, J. (1575). *Examen de los ingenios para las sciencias*. English translation by Richard Carew "The examination of men's wits", new edition by Rocío G. Sumillera, Modern Humanities Research Association, 2014.

Humies (2004). Humies awards. www.human-competitive.org, http://www.genetic-programming.org/combined.php.

Humphrey, N. K. (1976). The social function of intellect. In Bateson, P. and Hinde, R., editors, *Growing points in ethology*, pages 303–317. Cambridge University Press.

Humphry, S. M. (2011). The role of the unit in physics and psychometrics – and rejoinder. *Measurement*, 9(1):1–24, 64–70.

Hunt, T. (1928). The measurement of social intelligence. *Journal of Applied Psychology*, 12(3):317.

Hutter, M. (2002). The fastest and shortest algorithm for all well-defined problems. *International Journal of Foundations of Computer Science*, 13:431–443.

Hutter, M. (2005). *Universal artificial intelligence: Sequential decisions based on algorithmic probability*. Springer.

Hutter, M. (2007). Universal algorithmic intelligence: A mathematical top→down approach. In Goertzel, B. and Pennachin, C., editors, *Artificial general intelligence, cognitive technologies*, pages 227–290. Springer.

Hutter, M. (2012a). Can intelligence explode? *Journal of Consciousness Studies*, 19(1–2):143–166.

Hutter, M. (2012b). One decade of universal artificial intelligence. In *Theoretical foundations of artificial general intelligence*, pages 67–88. Springer.

Hutter, M., Legg, S., and Vitányi, P. M. (2007). Algorithmic probability. *Scholarpedia*, 2(8):2572.

Igel, C. and Toussaint, M. (2005). A no-free-lunch theorem for non-uniform distributions of target functions. *Journal of Mathematical Modelling and Algorithms*, 3(4):313–322.

Inoue, S. and Matsuzawa, T. (2007). Working memory of numerals in chimpanzees. *Current Biology*, 17(23):R1004–R1005.

Insa-Cabrera, J. (2016). *Towards a universal test of social intelligence*. PhD thesis, Universitat Politècnica de València.

Insa-Cabrera, J., Benacloch-Ayuso, J. L., and Hernández-Orallo, J. (2012a). On measuring social intelligence: Experiments on competition and cooperation. In Bach, J., Goertzel, B., and Iklé, M., editors, 5th International Conference on Artificial

General Intelligence, pages 126–135. *Lecture Notes in Computer Science*, volume 7716. Springer.

Insa-Cabrera, J., Dowe, D. L., España-Cubillo, S., Hernández-Lloreda, M. V., and Hernández-Orallo, J. (2011a). Comparing humans and AI agents. In Schmidhuber, J., Thórisson, K., and Looks, M., editors, 4th International Conference on *A*rtificial general intelligence, pages 122–132. *Lecture Notes in Artificial Intelligence*, volume 6830. Springer.

Insa-Cabrera, J., Dowe, D. L., and Hernández-Orallo, J. (2011b). Evaluating a reinforcement learning algorithm with a general intelligence test. In Lozano, J., Gamez, J., and Moreno, J., editors, *C*urrent Topics in Artificial Intelligence: CAEPIA 2011. *Lecture Notes in Artificial Intelligence*, volume 7023, pages 1–11. Springer.

Insa-Cabrera, J. and Hernández-Orallo, J. (2014). Definition and properties to assess multi-agent environments as social intelligence tests. CoRR arxiv preprint: abs/1408.6350.

Insa-Cabrera, J., Hernández-Orallo, J., Dowe, D., na, S. E., and Hernández-Lloreda, M. (2012b). The anYnt project intelligence test : Lambda – one. In Muller, V. and Ayesh, A., editors, *AISB/IACAP 2012 Symposium "Revisiting Turing and His Test"*, pages 20–27. Society for the Study of Artificial Intelligence and the Simulation of Behaviour.

Ishii, H., Ogura, M., Kurisu, S., Komura, A., Takanishi, A., Iida, N., and Kimura, H. (2006). Experimental study on task teaching to real rats through interaction with a robotic rat. In *From animals to animats 9*, pages 643–654. Springer.

Jaakkola, K., Fellner, W., Erb, L., Rodriguez, M., and Guarino, E. (2005). Understanding of the concept of numerically 'less' by bottlenose dolphins (tursiops truncatus). *Journal of Comparative Psychology*, 119(3):296.

Jacobs, G. H. and Nathans, J. (2009). The evolution of primate color vision. *Scientific American*, 300(4):56–63.

Jacoff, A., Messina, E., Weiss, B. A., Tadokoro, S., and Nakagawa, Y. (2003). Test arenas and performance metrics for urban search and rescue robots. In *Intelligent Robots and Systems, 2003 (IROS 2003)*, volume 4, pages 3396–3403. IEEE.

Janssen, R., Tuerlinckx, F., Meulders, M., and De Boeck, P. (2000). A hierarchical IRT model for criterion-referenced measurement. *Journal of Educational and Behavioral Statistics*, 25(3):285–306.

Japkowicz, N. and Shah, M. (2011). *Evaluating learning algorithms*. Cambridge University Press.

Jaulent, E. (2010). El Ars Generalis ultima de Ramon Llull: Presupuestos metafísicos y éticos. *Anales del Seminario de Historia de la Filosofía*, 27:87–113.

Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical Review*, 106(4):620.

Jelbert, S. A., Taylor, A. H., Cheke, L. G., Clayton, N. S., and Gray, R. D. (2014). Using the Aesop's fable paradigm to investigate causal understanding of water displacement by new Caledonian crows. *PloS ONE*, 9(3):e92895.

Jensen, A. R. (1980). Chronometric analysis of intelligence. *Journal of Social and Biological Structures*, 3(2):103–122.

Jensen, A. R. (1993). Why is reaction time correlated with psychometric g? *Current Directions in Psychological Science*, 2(2):53–56.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Praeger.

Jensen, A. R. (2000a). Artificial intelligence and g theory concern different phenomena. *Psycoloquy*, 11(86), Article 47.

Jensen, A. R. (2000b). A nihilistic philosophy of science for a scientific psychology? *Psycoloquy*, 11(88): Article 49.

Jensen, A. R. (2003). Regularities in Spearman's law of diminishing returns. *Intelligence*, 31(2):95–105.

John, O. P., Naumann, L. P., and Soto, C. J. (2008). Paradigm shift to the integrative big five trait taxonomy. In *Handbook of personality: Theory and research*, 3rd edition, pages 114–158. Guilford Press.

Johnson, M., Hofmann, K., Hutton, T., and Bignell, D. (2016). The Malmo platform for artificial intelligence experimentation. In *International Joint Conference on Artificial Intelligence (IJCAI)*, pages 4246–4247. IJCAI.

Johnson, N. L. and Kotz, S. (2011). Guttman, Louis in *Leading personalities in statistical sciences: from the seventeenth century to the present*, pp. 112–117, John Wiley.

Johnson, W. L. (1992). Needed: A new test of intelligence. *ACM SIGART Bulletin*, 3(4):7–9.

Johnson-Laird, P. N. (1986). *Mental models*. Harvard University Press.

Johnson-Laird, P. N., Khemlani, S. S., and Goodwin, G. P. (2015). Logic, probability, and human reasoning. *Trends in Cognitive Sciences*, 19(4):201–214.

Jolly, A. (1966). Lemur social behavior and primate intelligence. *Science*, 153(3735):501–506.

Jones, R. M. and Wray, R. E. (2011). Evaluating integrated, knowledge-rich cognitive systems. In *AAAI Fall Symposium: Advances in Cognitive Systems*, pages 138–145. Association for the Advancement of Artificial Intelligence.

Joosse, M., Lohse, M., Gallego-Pérez, J., and Evers, V. (2013). What you do is who you are: The role of task context in perceived social robot personality. In *2013 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2134–2139. IEEE.

Just, C. (2011). A review of literature on the general factor of personality. *Personality and Individual Differences*, 50(6):765–771.

Kak, S. (2000). Grading intelligence in machines: Lessons from animal intelligence. In Meystel, A. M. and Messina, E. R., editors, *Measuring the performance and intelligence of systems: Proceedings of the 2000 PerMIS Workshop*, pages 331–334. NIST Special Publication 970. NIST.

Kamar, E., Hacker, S., and Horvitz, E. (2012). Combining human and machine intelligence in large-scale crowdsourcing. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, volume 1, pages 467–474. International Foundation for Autonomous Agents and Multiagent Systems.

Kaminka, G. A., Frank, I., Arai, K., and Tanaka-Ishii, K. (2003). Performance competitions as research infrastructure: Large scale comparative studies of multi-agent teams. *Autonomous Agents and Multi-Agent Systems*, 7(1–2):121–144.

Kamphaus, R. W. (2005). *Clinical assessment of child and adolescent intelligence*. Springer Science and Business Media.

Kamphaus, R. W. and Frick, P. J. (2005). *Clinical assessment of child and adolescent personality and behavior*. Springer Science and Business Media.

Kanefsky, B. and Taylor, W. (1991). Where the really hard problems are. In *International Joint Conference on Artificial Intelligence (IJCAI)*, volume 91, pages 163–169. IJCAI.

Kant, I. (1785). *Grundlegung zur Metaphysik der Sitten*. Hartknoch, Riga, German-English edition "Groundwork of the Metaphysics of Morals", edited and translated by Mary Gregor and Jens Timmermann, 2011.

Karas, P. J., Mikell, C. B., Christian, E., Liker, M. A., and Sheth, S. A. (2013). Deep brain stimulation: A mechanistic and clinical update. *Neurosurgical Focus*, 35(5):E1.

Karpov, I. V., Schrum, J., and Miikkulainen, R. (2012). Believable bot navigation via playback of human traces. In Hingston, P., editor, *Believable bots*, pages 151–170. Springer.

Keedwell, E. (2010). Towards a staged developmental intelligence test for machines. In *Towards a Comprehensive Intelligence Test (TCIT): Reconsidering the Turing Test for the 21st Century Symposium*, pages 28–32. Society for the Study of Artificial Intelligence and Simulation of Behaviour.

Keith, T. Z. and Reynolds, M. R. (2010). Cattell–Horn–Carroll abilities and cognitive tests: What we've learned from 20 years of research. *Psychology in the Schools*, 47(7):635–650.

Kelley, T. L. (1927). *Interpretation of educational measurements.* World Book.

Kelley, T. L. (1928). *Crossroads in the mind of man: A study of differentiable mental abilities.* Stanford University Press.

Kemper, S. and Sumner, A. (2001). The structure of verbal abilities in young and older adults. *Psychology and Aging*, 16(2):312.

Ketter, W. and Symeonidis, A. (2012). Competitive benchmarking: Lessons learned from the trading agent competition. *AI Magazine*, 33(2):103.

Keynes, J. M. (1931). *Essays in persuasion*. Macmillan.

Kihlstrom, J. F. and Cantor, N. (2000). Social intelligence. In Sternberg, R., editor, *Handbook of intelligence*, pages 359–379. Cambridge University Press.

Kim, J. (2006). Task difficulty as a predictor and indicator of web searching interaction. In *CHI'06 Extended Abstracts on Human Factors in Computing Systems*, pages 959–964. ACM.

Kim, J. H., Kim, D. H., Kim, Y. J., and Seow, K. T. (2004). *Soccer robotics*. Springer.

King, J. E. and Figueredo, A. J. (1997). The five-factor model plus dominance in chimpanzee personality. *Journal of Research in Personality*, 31(2):257–271.

Kirsh, D. (1991). Today the earwig, tomorrow man? *Artificial Intelligence*, 47(1):161–184.

Kitano, H., Asada, M., Kuniyoshi, Y., Noda, I., and Osawa, E. (1997). Robocup: The robot world cup initiative. In *International Conference on Autonomous Agents*, pages 340–347. International Foundation for Autonomous Agents and Multiagent Systems.

Kittur, A., Nickerson, J. V., Bernstein, M., Gerber, E., Shaw, A., Zimmerman, J., Lease, M., and Horton, J. (2013). The future of crowd work. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, pages 1301–1318. ACM.

Klein, B. A., Stein, J., and Taylor, R. C. (2012). Robots in the service of animal behavior. *Communicative and Integrative Biology*, 5(5):466–472.

Kleiner, K. (2011). Who are you calling bird-brained? An attempt is being made to devise a universal intelligence test. *The Economist*, 398(8723):82.

Klenk, M., Forbus, K., Tomai, E., and Kim, H. (2011). Using analogical model formulation with sketches to solve Bennett mechanical comprehension test problems. *Journal of Experimental and Theoretical Artificial Intelligence*, 23(3):299–327.

Kline, P. (1998). *The new psychometrics: Science, psychology and measurement*. Routledge.

Kline, P. (2000). *Handbook of psychological testing*, 2nd edition. Routledge.

Knuth, D. E. (1973). *Sorting and searching*, volume 3 of *The art of computer programming*. Addison-Wesley.

Koehler, O. (1941). Vom Erlernen unbenannter Anzahlen bei Vögeln. *Naturwissenschaften*, 29(14):201–218.

Koene, R. A. (2012). Embracing competitive balance: The case for substrate-independent minds and whole brain emulation. In Eden, A. H., Moor, J. H., Soraker, J. H., and Steinhart, E., editors, *Singularity hypotheses*, pages 241–267. Springer.

Korsgaard, C. M. (1996). *Creating the kingdom of ends*. Cambridge University Press.

Kosinski, M., Bachrach, Y., Kasneci, G., Van-Gael, J., and Graepel, T. (2012). Crowd IQ: Measuring the intelligence of crowdsourcing platforms. In *Proceedings of the 4th Annual ACM Web Science Conference*, pages 151–160. ACM.

Kosinski, M., Stillwell, D., and Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 110(15):5802–5805.

Kowalski, D., editor (2012). *The big bang theory and philosophy: Rock, paper, scissors, Aristotle, Locke*. John Wiley.

Koza, J. R. (2010). Human-competitive results produced by genetic programming. *Genetic Programming and Evolvable Machines*, 11(3–4):251–284.

Krause, J., Winfield, A. F., and Deneubourg, J.-L. (2011a). Interactive robots in experimental biology. *Trends in Ecology and Evolution*, 26(7):369–375.

Krause, S., James, R., Faria, J. J., Ruxton, G. D., and Krause, J. (2011b). Swarm intelligence in humans: diversity can trump ability. *Animal Behaviour*, 81(5):941–948.

Krueger, J. and Osherson, D. (1980). On the psychology of structural simplicity. In Jusczyk, P. W. and Klein, R. M., editors, *The nature of thought: Essays in honor of D. O. Hebb*, pages 187–205. Psychology Press.

Kubinyi, E., Virányi, Z., and Miklósi, Á. (2007). Comparative social cognition: From wolf and dog to humans. *Comparative Cognition and Behavior Reviews*, 2:26– 46.

Kuhlmann, F. (1939). *Tests of mental development: A complete scale for individual examination*. Educational Test Bureau, Educational Publishers.

Kuncheva, L. I. (2004). *Combining pattern classifiers: Methods and algorithms*. John Wiley.

Kurzweil, R. (2005). *The singularity is near: When humans transcend biology*. Penguin.

Lacroix, B., Mathieu, P., and Kemeny, A. (2013). Formalizing the construction of populations in multi-agent simulations. *Engineering Applications of Artificial Intelligence*, 26(1):211–226.

Laird, J. E. (2012). *The Soar cognitive architecture*. MIT Press.

Laird, J. E., Newell, A., and Rosenbloom, P. S. (1987). Soar: An architecture for general intelligence. *Artificial Intelligence*, 33(1):1–64.

Laird, J. E. and Wray, R. E. (2010). Cognitive architecture requirements for achieving AGI. In Hutter, M., Baum, E., and Kitzelmann, E., editors, Proceedings of the *3rd*

International Conference AGI, Artificial general intelligence, pages 79–84. Advances in Intelligent Systems Research. Atlantis Press.

Laird, J. E., Wray, R. E., Marinier, R., and Langley, P. (2009). Claims and challenges in evaluating human-level intelligent systems. In *the Second Conference on Artificial General Intelligence*, pages 91–96. Atlantis Press.

Lake, B. M., Salakhutdinov, R., and Tenenbaum, J. B. (2015). Human-level concept learning through probabilistic program induction. *Science*, 350(6266):1332– 1338.

Lambiotte, R. and Kosinski, M. (2014). Tracking the digital footprints of personality. *Proceedings of the IEEE*, 102(12):1934–1939.

Langford, J. (2005). Clever methods of overfitting. *Machine Learning (Theory)*. http://hunch.net/?p=22.

Langley, P. (2011). Artificial intelligence and cognitive systems. *AISB Quarterly*, 133:1–4.

Lass, R. N., Sultanik, E. A., and Regli, W. C. (2009). Metrics for multiagent systems. In *Performance evaluation and benchmarking of intelligent systems*, pages 1–19. Springer.

Lasswitz, K. (1901). *The 'Universal Library'*. Originally in "Traumkristalle", English translation in "Fantasia mathematica", Clifton Fadiman, 1958, translated by Willy Ley, pp. 237–243.

Lattimore, T. (2013). *Theory of general reinforcement learning*. PhD thesis, Australian National University.

Lattimore, T. and Hutter, M. (2013). No free lunch versus Occam's razor in supervised learning. In Dowe, D. L., editor, *Algorithmic probability and friends: Bayesian prediction and artificial intelligence*, pages 223–235. *Lecture Notes in Computer Science*, volume 7070. Springer.

Laughlin, P. R. (2011). *Group problem solving*. Princeton University Press.

Law, E. and Von Ahn, L. (2011). Human computation. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 5(3):1–121.

Lee, S., Bang, W.-C., and Bien, Z.-Z. (2000). Measure of system intelligence: An engineering perspective. In Meystel, A. M. and Messina, E. R., editors, *Measuring the performance and intelligence of systems: Proceedings of the 2000 PerMIS Workshop*, pages 112–117. NIST Special Publication 970. NIST.

Leeuwenberg, E. L. (1969). Quantitative specification of information in sequential patterns. *Psychological Review*, 76(2):216.

Leeuwenberg, E. L. J. and Van Der Helm, P. A. (2012). *Structural information theory: The simplicity of visual form*. Cambridge University Press.

Legg, S. (2008). *Machine super intelligence*. PhD thesis, Department of Informatics, University of Lugano.

Legg, S. and Hutter, M. (2005). A universal measure of intelligence for artificial agents. In *International Joint Conference on Artificial Intelligence, IJCAI*, pages 1509–1510. IJCAI.

Legg, S. and Hutter, M. (2007a). A collection of definitions of intelligence. *Frontiers in Artificial Intelligence and Applications*, 157:17.

Legg, S. and Hutter, M. (2007b). Tests of machine intelligence. In Lungarella, M., Iida, F., Bongard, J., and Pfeifer, R., editors, *50 Years of Artificial Intelligence*, pages 232–242. *Lecture Notes in Computer Science*, volume 4850. Springer.

Legg, S. and Hutter, M. (2007c). Universal intelligence: A definition of machine intelligence. *Minds and Machines*, 17(4):391–444.

Legg, S. and Veness, J. (2013). An approximation of the universal intelligence measure. In Dowe, D. L., editor, *Algorithmic probability and friends: Bayesian prediction and artificial intelligence*, pages 236–249. *Lecture Notes in Computer Science*, volume 7070. Springer.

Leiber, J. (1985). *Can animals and machines be persons? A dialogue*. Hackett.

Leike, J. and Hutter, M. (2015). Bad universal priors and notions of optimality. *JMLR: Workshop and Conference Proceedings*, 40:1–16.

Lem, S. (1971). *The Futurological Congress (from the memoirs of Ijon Tichy)*. Translated into English in 1974 by Michael Kandel. Seabury Press.

Levesque, H. J. (2014). On our best behaviour. *Artificial Intelligence*, 212:27–35.

Levesque, H. J., Davis, E., and Morgenstern, L. (2012). The Winograd schema challenge. In *Thirteenth International Conference on the Principles of Knowledge Representation and Reasoning*, pages 552–561.

Levin, L. A. (1973). Universal sequential search problems. *Problems of Information Transmission*, 9(3):265–266.

Levin, L. A. (1974). Laws of information conservation (nongrowth) and aspects of the foundation of probability theory. *Problemy Peredachi Informatsii (Problems of Information Transmission)*, 10(3):30–35.

Levin, L. A. (1986). Average case complete problems. *SIAM Journal on Computing*, 15:285–286.

Levin, L. A. (1997). The century of bread and circus. *Tax Notes*, 77(7):858–860.

Levin, L. A. (2013). Universal heuristics: How do humans solve “unsolvable” problems? In Dowe, D. L., editor, *Algorithmic probability and friends: Bayesian prediction and artificial intelligence*, pages 53–54. *Lecture Notes in Computer Science*, volume 7070. Springer.

Levine, J., Congdon, C. B., Ebner, M., Kendall, G., Lucas, S. M., Miikkulainen, R., Schaul, T., Thompson, T., Lucas, S. M., Mateas, M., et al. (2013). General video game playing. *Artificial and Computational Intelligence in Games*, 6:77–83.

Levy, F. and Murnane, R. J. (2012). *The new division of labor: How computers are creating the next job market*. Princeton University Press.

Li, M. and Vitányi, P. (2008). *An introduction to Kolmogorov complexity and its applications* (3rd edition). Springer.

Li, Y. and Belkin, N. J. (2008). A faceted approach to conceptualizing tasks in information seeking. *Information Processing and Management*, 44(6):1822–1837.

Lillibridge, M. D., Abadi, M., Bharat, K., and Broder, A. Z. (2001). Method for selectively restricting access to computer systems. Filed on 13 April 1998 and granted on 27 February 2001, US Patent 6,195,698.

List, C. and Puppe, C. (2009). Judgment aggregation: A survey. In P. Anand, C. Puppe, P. Pattanaik (Eds.) *Rational and Social Choice*, pages 457–482. Oxford University Press.

Little, V. L. and Bailey, K. G. (1972). Potential intelligence or intelligence test potential? A question of empirical validity. *Journal of Consulting and Clinical Psychology*, 39(1):168.

Liu, J., Abbass, H. A., Green, D. G., and Zhong, W. (2012). Motif difficulty (MD): A predictive measure of problem difficulty for evolutionary algorithms using network motifs. *Evolutionary Computation*, 20(3):321–347.

Liu, J., Gwizdka, J., Liu, C., and Belkin, N. J. (2010). Predicting task difficulty for different task types. *Proceedings of the American Society for Information Science and Technology*, 47(1):1–10.

Liu, P. and Li, Z. (2012). Task complexity: A review and conceptualization framework. *International Journal of Industrial Ergonomics*, 42(6):553–568.

Livingston, S. A. (2006). Item analysis. In Downing, S. M. and Haladyna, T. M., editors, *Handbook of test development*, pages 421–441. Lawrence Erlbaum Associates.

Livingstone, D. (2006). Turing's test and believable AI in games. *Computers in Entertainment*, 4(1):6.

Llull, R. (1295). *Cant de Ramon. Lo desconhort*. Ensiola Editorial, 2006.

Llull, R. (1305). *Ars generalis ultima*. Edition by Alois Madre, ROL XIV. Turnhout.

Loebner, H. (2008). How to hold a Turing test contest. In Epstein, R., Roberts, G., and Beber, G., editors, *Parsing the Turing test: Philosophical and methodological issues in the quest for the thinking computer*, pages 173–179. Springer.

Loehlin, J. C. (2000). Group differences in intelligence. In Sternberg, R., editor, *Handbook of intelligence*, pages 176–193. Cambridge University Press.

Logie, R. H. (2014). *Visuo-spatial working memory*. Psychology Press.

Lohman, D. (2000). Complex information processing and intelligence. In Sternberg, R., editor, *Handbook of intelligence*, pages 285–340. Cambridge University Press.

Lohman, D. F., Lakin, J. M., Sternberg, R., and Kaufman, S. (2011). Reasoning and intelligence. In R. J. Sternberg, editor, *Handbook of intelligence*, pages 419–441. Cambridge University Press.

Long, D. and Fox, M. (2003). The 3rd international planning competition: Results and analysis. *Journal of Artificial Intelligence Research*, 20:1–59.

Loosemore, R. and Goertzel, B. (2012). Why an intelligence explosion is probable. In Eden, A. H., Moor, J. H., Soraker, J. H., and Steinhart, E., editors, *Singularity hypotheses*, pages 83–98. Springer.

Lord, F. M., Novick, M. R., and Birnbaum, A. (1968). *Statistical theories of mental test scores.* Addison-Wesley.

Lovett, A. and Forbus, K. (2011). Cultural commonalities and differences in spatial problem-solving: A computational analysis. *Cognition*, 121(2):281–287.

Lovett, A., Forbus, K., and Usher, J. (2010). A structure-mapping model of Raven's Progressive Matrices. In *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, volume 10, pages 2761–2766. Cognitive Science Society.

Lovett, A., Lockwood, K., and Forbus, K. (2008). A computational model of the visual oddity task. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, volume 25, page 29. Cognitive Science Society.

Lycan, W. G. (1985). Abortion and the civil rights of machines. In *Morality and universality*, pages 139–156. Springer.

Lyle, J. (2008). Collective problem solving: Are the many smarter than the few? *Durhan Anthropology Journal*, 15(1):23–58.

Lynn, R. (1996). *Dysgenics: Genetic deterioration in modern populations*. Praeger.

Lynn, R. (2001). *Eugenics: A reassessment*. Greenwood.

Lynn, R. and Harvey, J. (2008). The decline of the world's IQ. *Intelligence*, 36(2):112–120.

Lyon, C., Nehaniv, C. L., and Saunders, J. (2012). Interactive language learning by robots: The transition from babbling to word forms. *PloS ONE*, 7(6):e38236.

Machluf, K., Liddle, J. R., and Bjorklund, D. F. (2013). An introduction to evolutionary developmental psychology. *Evolutionary Psychology*, 12(2):264–272.

Macià, N. and Bernadó-Mansilla, E. (2014). Towards UCI+: A mindful repository design. *Information Sciences*, 261:237–262.

MacInnes, W. J., Armstrong, B. C., Pare, D., Cree, G. S., and Joordens, S. (2009). Everyone's a critic: Memory models and uses for an artificial Turing judge. In the 2nd *Conference on Artificial General Intelligence*, pages 132–137. Atlantis Press.

Mackintosh, N. and Mackintosh, N. J. (2011). *IQ and human intelligence*. Oxford University Press.

Madden, M. G. and Howley, T. (2004). Transfer of experience between reinforcement learning environments with progressive difficulty. *Artificial Intelligence Review*, 21(3–4):375–398.

Madhavan, R., Tunstel, E., and Messina, E. (2009). *Performance evaluation and benchmarking of intelligent systems*. Springer.

Madsen, P. T., Jensen, F. H., Carder, D., and Ridgway, S. (2012). Dolphin whistles: A functional misnomer revealed by heliox breathing. *Biology Letters*, 8(2):211– 213.

Mahoney, M. V. (1999). Text compression as a test for artificial intelligence. In *Proceedings of the National Conference on Artificial Intelligence*, pages 970–970. AAAI.

Mahrer, A. R. (1958). Potential intelligence: A learning theory approach to description and clinical implication. *Journal of General Psychology*, 59(1):59–71.

Malan, K. M. and Engelbrecht, A. P. (2013). A survey of techniques for characterising fitness landscapes and some possible ways forward. *Information Sciences*, 241:148–163.

Malone, T. W. and Bernstein, M. S., editors (2015). *Handbook of collective intelligence*. MIT Press.

Mandolesi, L., Addona, F., Foti, F., Menghini, D., Petrosini, L., and Vicari, S. (2009). Spatial competences in Williams syndrome: A radial arm maze study. *International Journal of Developmental Neuroscience*, 27(3):205–213.

Mar, H. (1996). *Psychological evaluation of children who are deaf-blind: An overview with recommendations for practice*. National Consortium on Deaf-Blindness.

Marché, C. and Zantema, H. (2007). The termination competition. In *Term rewriting and applications*, pages 303–313. Springer.

Marcolino, L. S., Xu, H., Jiang, A. X., Tambe, M., and Bowring, E. (2014). Give a hard problem to a diverse team: Exploring large action spaces. *Proceedings of 28th AAAI*. AAAI.

Marcus, G., Rossi, F., and Veloso, M. (2015). Beyond the Turing test. AAAI workshop, http://www.math.unipd.it/~frossi/BeyondTuring2015/.

Marcus, G., Rossi, F., and Veloso, M., eds. (2016). Beyond the Turing test (special issue). *AI Magazine*, 37(1):3–101.

Markowetz, A., Błaszkiewicz, K., Montag, C., Switala, C., and Schlaepfer, T. E. (2014). Psycho-informatics: Big data shaping modern psychometrics. *Medical Hypotheses*, 82(4):405–411.

Marshalek, B., Lohman, D. F., and Snow, R. E. (1983). The complexity continuum in the radex and hierarchical models of intelligence. *Intelligence*, 7(2):107–127.

Martínez-Plumed, Hernández-Orallo, J., Ramírez-Quintana, M., and Ferri, C. (2015). A computational analysis of general intelligence tests for evaluating cognitive development. Unpublished manuscript.

Martínez-Plumed, F., Ferri, C., Hernández-Orallo, J., and Ramírez-Quintana, M. J. (2015). Knowledge acquisition with forgetting: An incremental and developmental setting. *Adaptive Behavior*, 23(5):283–299.

Masum, H. and Christensen, S. (2003). The Turing ratio: A framework for open-ended task metrics. *Journal of Evolution and Technology*, 13(2).

Masum, H., Christensen, S., and Oppacher, F. (2002). The Turing ratio: Metrics for open-ended tasks. In *GECCO, Proceedings of the Genetic and Evolutionary Computation Conference*, pages 973–980. Morgan Kaufmann.

Matellán-Olivera, V. et al. (2011). Comparative cognition: Animals and robots. *ERCIM News*, 84:31–32.

Maturana, H. R. and Varela, F. J. (1980). *Autopoiesis and cognition: The realization of the living*. Kluwer.

Matzel, L. D. and Kolata, S. (2010). Selective attention, working memory, and animal intelligence. *Neuroscience and Biobehavioral Reviews*, 34(1):23–30.

Mayfield, J. E. (2007). Minimal history, a theory of plausible explanation. *Complexity*, 12(4):48–53.

Mayfield, J. E. (2013). *The engine of complexity: Evolution as computation*. Columbia University Press.

Mayr, E. (1982). *The growth of biological thought: Diversity, evolution, and inheritance*. Harvard University Press.

McCarthy, J. (1979). Ascribing mental qualities to machines. In Ringle, M., editor, *Philosophical perspectives in artificial intelligence*, pages 161–195. Humanities Press.

McCarthy, J. (1983). Artificial intelligence needs more emphasis on basic research: President's quarterly message. *AI Magazine*, 4(4):5.

McCarthy, J. (1987). Generality in artificial intelligence. *Communications of the ACM*, 30(12):1030–1035.

McCarthy, J. (2007). What is artificial intelligence? Technical report, Stanford University. http://www-formal.stanford.edu/jmc/whatisai.html.

McCarthy, J. (2008). The well-designed child. *Artificial Intelligence*, 172(18):2003–2014.

McClelland, D. C. (1973). Testing for competence rather than for "intelligence". *American Psychologist*, 28(1):1.

McCorduck, P. (2004). *Machines who think*. A K Peters/CRC Press.

McCrae, R. and Costa, P. (2010). *Neo inventories professional manual*. Psychological Assessment Resources.

McDermott, D. (1997). How intelligent is Deep Blue? *New York Times*, May 14.

McDermott, D. (2007). Level-headed. *Artificial Intelligence*, 171(18):1183–1186.

McDermott, D. (2010). Don't improve the Turing test, abandon it. In Ayesh, A., Bishop, M., Floridi, L., and Warwick, K., editors, *AISB 2010 Symposium "Towards a Comprehensive Intelligence Test (TCIT): Reconsidering the Turing Test for the 21st Century"*, pages 37–41. Society for the Study of Artificial Intelligence and the Simulation of Behaviour.

McDermott, J., White, D. R., Luke, S., Manzoni, L., Castelli, M., Vanneschi, L., Jaśkowski, W., Krawiec, K., Harper, R., Jong, K. D., and O'Reilly, U.-M. (2012). Genetic programming needs better benchmarks. In *International Conference on Genetic and Evolutionary Computation*, pages 791–798. ACM.

McGrath, J. E. (1984). *Groups: Interaction and performance*. Prentice Hall.

McIlwain, K. L., Merriweather, M. Y., Yuva-Paylor, L. A., and Paylor, R. (2001). The use of behavioral test batteries: Effects of training history. *Physiology and Behavior*, 73(5):705–717.

Melkikh, A. V. (2014). The no free lunch theorem and hypothesis of instinctive animal behavior. *Artificial Intelligence Research*, 3(4):43.

Mellenbergh, G. J. (1994). Generalized linear item response theory. *Psychological Bulletin*, 115(2):300.

Menary, R. (2010). *The extended mind*. MIT Press.

Merrell, K. W. (2001). Assessment of children's social skills: Recent developments, best practices, and new directions. *Exceptionality*, 9(1–2):3–18.

Mertens, S. (1998). Phase transition in the number partitioning problem. *Physical Review Letters*, 81(20):4281.

Messina, E., Meystel, A., and Reeker, L. (2001). PerMIS 2001, white paper. In Meystel, A. M. and Messina, E. R., editors, *Measuring the performance and intelligence of systems: Proceedings of the 2001 PerMIS Workshop*, pages 3–15. NIST Special Publication 982. NIST.

Meystel, A. (2000a). Appendix A. The preliminary discussions. In Meystel, A. M. and Messina, E. R., editors, *Measuring the performance and intelligence of systems: Proceedings of the 2000 PerMIS Workshop*, pages 563–642. NIST Special Publication 970. NIST.

Meystel, A. (2000b). Permis 2000 white paper: Measuring performance and intelligence of systems with autonomy. In Meystel, A. M. and Messina, E. R., editors, *Measuring the performance and intelligence of systems: Proceedings of the 2000 PerMIS Workshop*, pages 1–34. NIST Special Publication 970. NIST.

Meystel, A., Albus, J., Messina, E., and Leedom, D. (2003a). Performance measures for intelligent systems: Measures of technology readiness. Technical report, DTIC Document.

Meystel, A., Albus, J., Messina, E., and Leedom, D. (2003b). Permis 2003 white paper: Performance measures for intelligent systems – measures of technology readiness. In Meystel, A. M. and Messina, E. R., editors, *Measuring the performance and intelligence of systems: Proceedings of the 2003 PerMIS Workshop*. NIST Special Publication 1014. NIST.

Meystel, A. and Andrusenko, J. (2001). Evaluating the performance of e-coli with genetic learning from simulated testing. In Meystel, A. M. and Messina, E. R., editors, *Measuring the performance and intelligence of systems: Proceedings of the 2001 PerMIS Workshop*. NIST Special Publication 982. NIST.

Michel, O. and Rohrer, F. (2008). The rat's life benchmark: Competing cognitive robots. In Madhavan, R. and Messina, E., editors, *Proceedings of the 8th Workshop on Performance Metrics for Intelligent Systems*, pages 43–49. ACM.

Michell, J. (2000). Normal science, pathological science and psychometrics. *Theory and Psychology*, 10(5):639–667.

Michell, J. (2008). Is psychometrics pathological science? *Measurement*, 6(1–2):7–24.

Michelucci, P. and Dickinson, J. L. (2016). The power of crowds. *Science*, 351(6268):32–33.

Miller, G. (2010). Social savvy boosts the collective intelligence of groups. *Science*, 330(6000):22–22.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2):81.

Miller, G. A. (1995). Wordnet: A lexical database for English. *Communications of the ACM*, 38(11):39–41.

Minsky, M. L., editor (1968). *Semantic information processing*. MIT Press.

Minsky, M. L. (1988). *Society of mind*. Simon and Schuster.

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., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., and Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(26 February):529–533.

Monroe, P. E. (1931). *Conference on examinations*. Teachers College, Columbia University.

Moor, J. (2003). *The Turing test: The elusive standard of artificial intelligence*, volume 30. Springer Science and Business Media.

Moore, C. and Mertens, S. (2011). *The nature of computation*. Oxford University Press.

Moravec, H. P. (1988). *Mind children: The future of robot and human intelligence*. Harvard University Press.

Morgan, C. L. (1903). *An introduction to comparative psychology*. W. Scott.

Morgenstern, L., Davis, E., and Ortiz-Jr., C. L. (2016). Planning, executing, and evaluating the Winograd schema challenge. *AI* Magazine, 37(1):50–54.

Morgenstern, L. and Ortiz, C. (2015). The Winograd schema challenge: Evaluating progress in commonsense reasoning. In *The Twenty-Ninth AAAI Conference on Artificial Intelligence*, pages 4024–4026. AAAI.

Morris, D. (1967). *The naked ape: A zoologist's study of the human animal*. Jonathan Cape.

Moss, F., Hunt, T., Omwake, K., and Ronning, M. (1925). *George Washington University social intelligence test*. Cen.

Mueller, E. T. (1987). Daydreaming and computation. Technical report, TR, CSD-870017. PhD dissertation, University of California, Los Angeles.

Mueller, S. T. (2008). Is the Turing Test still relevant? A plan for developing the cognitive decathlon to test intelligent embodied behavior. In *19th Midwest Artificial Intelligence and Cognitive Science Conference*.

Mueller, S. T. (2010). A partial implementation of the BICA cognitive decathlon using the psychology experiment building language (PEBL). *International Journal of Machine Consciousness*, 2(02):273–288.

Mueller, S. T., Jones, M., Minnery, B., and Hiland, J. M. (2007). The BICA cognitive decathlon: A test suite for biologically-inspired cognitive agents. In *Proceedings of Behavior Representation in Modeling and Simulation Conference, Norfolk*.

Mueller, S. T. and Piper, B. J. (2014). The Psychology Experiment Building Language (PEBL) and PEBL test battery. *Journal of Neuroscience Methods*, 222:250– 259.

Murphy, K. R. (2006). *A critique of emotional intelligence: What are the problems and how can they be fixed?* Psychology Press.

Murray, A. L., Dixon, H., and Johnson, W. (2013). Spearman's law of diminishing returns: A statistical artifact? *Intelligence*, 41(5):439–451.

Myers, G. J., Sandler, C., and Badgett, T. (2011). *The art of software testing*. John Wiley.

Nakagaki, T., Yamada, H., and Tóth, Á. (2000). Intelligence: Maze-solving by an amoeboid organism. *Nature*, 407(6803):470–470.

Naor, M. (1996). Verification of a human in the loop or identification via the Turing test. http://www.wisdom.weizmann.ac.il/~naor/PAPERS/human_abs.html.

Narr, K. L., Woods, R. P., Thompson, P. M., Szeszko, P., Robinson, D., Dimtcheva, T., Gurbani, M., Toga, A. W., and Bilder, R. M. (2007). Relationships between IQ and regional cortical gray matter thickness in healthy adults. *Cerebral Cortex*, 17(9):2163–2171.

Nass, C., Moon, Y., Fogg, B., Reeves, B., and Dryer, D. C. (1995). Can computer personalities be human personalities? *International Journal of Human-Computer Studies*, 43(2):223–239.

Neary, T. and Woods, D. (2006). P-completeness of cellular automaton rule 110. In *Automata, Languages and Programming*, pages 132–143. Springer.

Nehaniv, C. L. and Dautenhahn, K. (2007). *Imitation and social learning in robots, humans and animals: Behavioural, social and communicative dimensions*. Cambridge University Press.

Nettle, D. (2006). The evolution of personality variation in humans and other animals. *American Psychologist*, 61(6):622.

Neumann, F., Reichenberger, A., and Ziegler, M. (2009). Variations of the Turing test in the age of Internet and virtual reality. In *KI 2009: Advances in Artificial Intelligence*, pages 355–362. Springer.

Newborg, J., Stock, J. R., Wnek, L., Guidubaldi, J., Svinicki, J., et al. (1984). *Battelle developmental inventory*. DLM Teaching Resources.

Newcomer, P. L. and Hammill, D. D. (2008). *Test of Language Development: Primary (TOLD-P: 4)*. PRO-ED.

Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In Chase, W., editor, *Visual information processing*, pages 283–308. Academic Press.

Newell, A. (1980). Physical symbol systems. *Cognitive Science*, 4(2):135–183.

Newell, A. (1990). *Unified theories of cognition*. Harvard University Press.

Newell, A., Shaw, J. C., and Simon, H. A. (1959). Report on a general problem-solving program. In International Conference on Information Processing, pages 256–264. UNESCO.

Newell, A. and Simon, H. A. (1961). GPS, a program that simulates human thought. Technical report, DTIC Document.

Newell, A. and Simon, H. A. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the ACM*, 19(3):113–126.

Ng, A. Y., Russell, S. J., et al. (2000). Algorithms for inverse reinforcement learning. In *International Conference on Machine Learning*, pages 663–670. Morgan Kaufman.

Nilsson, N. J. (1980). *Principles of artificial intelligence*. Morgan Kaufmann.

Nilsson, N. J. (1995). Eye on the prize. *AI Magazine*, 16(2):9.

Nilsson, N. J. (2006). Human-level artificial intelligence? Be serious! *AI Magazine*, 26(4):68.

Nilsson, N. J. (2009). *The quest for artificial intelligence*. Cambridge University Press.

Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D. F., and Turkheimer, E. (2012). Intelligence: New findings and theoretical developments. *American Psychologist*, 67(2):130.

Nivel, E., Thórisson, K. R., Steunebrink, B. R., Dindo, H., Pezzulo, G., Rodriguez, M., Hernandez, C., Ognibene, D., Schmidhuber, J., Sanz, R., Helgason, H. P., Chella, A., and Jonsson, G. K. (2014). Autonomous acquisition of natural language. In A. P. dos Reis, P. Kommers, and P. Isaías (Eds.), *Proceedings of the IADIS International Conference on Intelligent Systems & Agents 2014* (ISA-14), pages 58–66.

Oates, T. and Chong, W. (2006). Book review: Marcus Hutter, universal artificial intelligence. *Artificial Intelligence*, 170:1222–1226.

Ohlsson, S., Sloan, R. H., Turán, G., and Urasky, A. (2013). Verbal IQ of a four-year-old achieved by an AI system. In *COMMONSENSE 2013, 11th International Symposium on Logical Formalizations of Commonsense Reasoning, Ayia Napa, Cyprus*, page 6. Commonsense Reading.

Ohlsson, S., Sloan, R. H., Turán, G., and Urasky, A. (2015). Measuring an artificial intelligence system's performance on a verbal IQ test for young children. CoRR arXiv preprint:1509.03390.

Olton, D. S. and Samuelson, R. J. (1976). Remembrance of places passed: Spatial memory in rats. *Journal of Experimental Psychology: Animal Behavior Processes*, 2(2):97.

Omohundro, S. (2012). Rational artificial intelligence for the greater good. In Eden, A. H., Moor, J. H., Soraker, J. H., and Steinhart, E., editors, *Singularity hypotheses*, pages 161–179. Springer.

Oppy, G. and Dowe, D. L. (2011). The Turing test. In Zalta, E. N., editor, *Stanford encyclopedia of philosophy*. Stanford University. http://plato.stanford.edu/entries/turing-test/.

Orponen, P., Ko, K.-I., Schöning, U., and Watanabe, O. (1994). Instance complexity. *Journal of the ACM*, 41(1):96–121.

Orseau, L. (2013). Asymptotic non-learnability of universal agents with computable horizon functions. *Theoretical Computer Science*, 473:149–156.

Orseau, L. (2014). Universal knowledge-seeking agents. *Theoretical Computer Science*, 519:127–139.

Orseau, L., Lattimore, T., and Hutter, M. (2013). Universal knowledge-seeking agents for stochastic environments. In *Algorithmic learning theory*, pages 158–172. Springer.

Orseau, L. and Ring, M. (2011). Self-modification and mortality in artificial agents. In J. Schmidhuber, K. R. Thórisson, and M. Looks, editors, *4th Conference on Artificial General Intelligence*, pages 1–10. *Lecture Notes in Artificial Intelligence*, volume 6830. Springer.

Ortigosa, A., Carro, R. M., and Quiroga, J. I. (2013). Predicting user personality by mining social interactions in facebook. *Journal of Computer and System Sciences*, 80:57–71.

Ortiz, C. L., Jr. (2016). Why we need a physically embodied Turing test and what it might look like. *AI* Magazine, 37(1):55–62.

Orwell, G. (1949). *Nineteen Eighty-Four*. Secker and Warburg.

O'Sullivan, M., Guilford, J., and deMille, R. (1965). *Measurement of social intelligence*. ERIC, University of South California.

Owens, R. E. (2015). *Language development: An introduction*, 9th edition. Pearson.

Page, S. E. (2008). *The difference: How the power of diversity creates better groups, firms, schools, and societies*. Princeton University Press.

Panait, L. and Luke, S. (2005). Cooperative multi-agent learning: The state of the art. *Autonomous Agents and Multi-Agent Systems*, 11(3):387–434.

Park, G., Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Kosinski, M., Stillwell, D. J., Ungar, L. H., and Seligman, M. E. P. (2015). Automatic personality assessment through social media language. *Journal of Personality and Social Psychology*, 108(6):934.

Park, H.-J., Kim, B. K., and Lim, K. Y. (2001). Measuring the machine intelligence quotient (MIQ) of human-machine cooperative systems. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 31(2):89–96.

Parker, S. T. and McKinney, M. L. (2012). *Origins of intelligence: The evolution of cognitive development in monkeys, apes, and humans*. Johns Hopkins University Press.

Pearce, D. (2004). *Hedonistic imperative*. http://www.hedweb.com/.

Penn, D. C., Holyoak, K. J., and Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences*, 31(2):109–130.

Perdue, B. M., Talbot, C. F., Stone, A. M., and Beran, M. J. (2012). Putting the elephant back in the herd: Elephant relative quantity judgments match those of other species. *Animal Cognition*, 15(5):955–961.

Perez, D., Samothrakis, S., Togelius, J., Schaul, T., Lucas, S., Couëtoux, A., Lee, J., Lim, C.-U., and Thompson, T. (2015). The 2014 general video game playing competition. *IEEE Transactions on Computational Intelligence and AI in Games*.

Persson, E. (2012). The moral status of extraterrestrial life. *Astrobiology*, 12(10):976–984.

Pervin, L. A. (1970). *Personality: Theory, assessment, and research.* John Wiley.

PETRL (2015). People for the ethical treatment of reinforcement learners. http://petrl.org/.

Piaget, J. (1936). *La naissance de l'intelligence chez l'enfant*. Delachaux et Niestlé. (English translation 1952: "The origins of intelligence in children").

Piaget, J. (1964). Cognitive development in children. *Journal of Research in Science Teaching*, 2(3):176–186.

Pinker, S. (1995). *The language instinct: The new science of language and mind*, volume 7529. Penguin.

Pinker, S. (2003). *The blank slate: The modern denial of human nature*. Penguin.

Pinker, S. (2009). *Language learnability and language development, with new commentary by the author*. Harvard University Press.

Pintner, R. (1921). Intelligence and its measurement: A symposium. *Journal of Educational Psychology*, 12(3):139.

Plasencia, A. (2017). *Is the universe a hologram? And other conversations with leading scientists*. MIT Press.

Plomin, R. (2001). The genetics of g in human and mouse. *Nature Reviews Neuroscience*, 2(2):136–141.

Plomin, R. and Asbury, K. (2005). Nature and nurture: Genetic and environmental influences on behavior. *Annals of the American Academy of Political and Social Science*, 600(1):86–98.

Plomin, R. and Deary, I. (2015). Genetics and intelligence differences: Five special findings. *Molecular Psychiatry*, 20:98–108.

Plomin, R., DeFries, J. C., Knopik, V. S., and Neiderheiser, J. (2013). *Behavioral Genetics*. Palgrave Macmillan.

Poland, J. and Hutter, M. (2005). Asymptotics of discrete MDL for online prediction. *IEEE Transactions on Information Theory*, 51(11):3780–3795.

Poland, J. and Hutter, M. (2006). MDL convergence speed for Bernoulli sequences. *Statistics and Computing*, 16(2):161–175.

Pollack, M. E., Brown, L., Colbry, D., McCarthy, C. E., Orosz, C., Peintner, B., Ramakrishnan, S., and Tsamardinos, I. (2003). Autominder: An intelligent cognitive orthotic system for people with memory impairment. *Robotics and Autonomous Systems*, 44(3):273–282.

Potthast, M., Hagen, M., Gollub, T., Tippmann, M., Kiesel, J., Rosso, P., Stamatatos, E., and Stein, B. (2013). Overview of the 5th international competition on plagiarism detection. In *CLEF 2013 Evaluation Labs and Workshop Working Notes Papers, 23–26 September, Valencia, Spain*.

Preiss, D. D. and Sternberg, R. J. (2005). Technologies for working intelligence. In Preiss, D. D. and Sternberg, R. J., editors, *Intelligence and technology: The impact of tools on the nature and development of human abilities*, pages 183–208. Routledge.

Preston, B. (1991). AI, anthropocentrism, and the evolution of 'intelligence'. *Minds and Machines*, 1(3):259–277.

Princeton University (2010). About Wordnet. http://wordnet.princeton.edu/.

Proudfoot, D. (2011). Anthropomorphism and AI: Turing's much misunderstood imitation game. *Artificial Intelligence*, 175(5):950–957.

Pullum, G. K. and Rogers, J. (2006). *Animal pattern-learning experiments: Some mathematical background*. Radcliffe Institute for Advanced Study/Harvard University.

Pullum, G. K. and Scholz, B. C. (2002). Empirical assessment of stimulus poverty arguments. *The Linguistic Review*, 18(1–2):9–50.

Pullum, G. K. and Scholz, B. C. (2009). For universals (but not finite-state learning) visit the zoo. *Behavioral and Brain Sciences*, 32(5):466–467.

Purdy, T., Grutter, K., Davanco, M., Srinivasan, K., and Taylor, J. (2016). Optomechanical quantum correlation thermometry. *Bulletin of the American Physical Society*. APS March Meeting 2016, abstract #P52.011. APS.

Quercia, D., Kosinski, M., Stillwell, D., and Crowcroft, J. (2011). Our Twitter profiles, our selves: Predicting personality with Twitter. In *Third International Conference on Social Computing (SOCIALCOM)*, pages 180–185. IEEE.

Quinn, A. J. and Bederson, B. B. (2011). Human computation: A survey and taxonomy of a growing field. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1403–1412. ACM.

Quiroga, R. Q. (2012). *Borges and memory: Encounters with the human brain*. MIT Press.

Ragni, M. and Neubert, S. (2012). Solving Raven's IQ-tests: An AI and cognitive modeling approach. In *ECAI*, pages 666–671. IOS Press.

Ragni, M. and Neubert, S. (2014). Analyzing Raven's intelligence test: Cognitive model, demand, and complexity. In Prade, H. and Richard, G., editors, *Computational approaches to analogical reasoning: Current trends*, pages 351–370. *Studies in Computational Intelligence*, volume 548. Springer.

Rajani, S. (2011). Artificial intelligence – man or machine. *International Journal of Information Technology*, 4(1):173–176.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainnment tests*. Danmarks paedagogiske Institut.

Rauscher, F. H., Shaw, G. L., and Ky, K. N. (1993). Music and spatial task performance. *Nature*, 365(14 October):611.

Raven, J. C., Court, J. H., and Raven, J. (1992). *Manual for Raven's Progressive Matrices and Vocabulary Scale*. Psychological Corporation.

Reader, S. M., Hager, Y., and Laland, K. N. (2011). The evolution of primate general and cultural intelligence. *Philosophical Transactions of the Royal Society, Series B*, 366(1567):1017–1027.

Regan, T. (1987). *The case for animal rights*. Springer.

Revelle, W. and Wilt, J. (2013). The general factor of personality: A general critique. *Journal of Research in Personality*, 47(5):493–504.

Reznikova, Z. (2007). *Animal intelligence: From individual to social cognition*. Cambridge University Press.

Reznikova, Z. and Ryabko, B. (2003). In the shadow of the binary tree: Of ants and bits. In *Proceedings of the 2nd Intl Workshop of the Mathematics and Algorithms of Social Insects*, pages 139–145. Georgian Institute of Technology.

Richardson, M. and Stalnaker, J. M. (1933). A note on the use of bi-serial r in test research. *Journal of General Psychology*, 8(2):463–465.

Riggio, R. E. (1986). Assessment of basic social skills. *Journal of Personality and Social Psychology*, 51(3):649.

Ring, M. and Orseau, L. (2011). Delusion, survival, and intelligent agents. In J. Schmidhuber, K. R. Thórisson, and M. Looks, editors, *4th Conference on Artificial General Intelligence*, pages 11–20. *Lecture Notes in Artificial Intelligence*, volume 6830. Springer.

Rissanen, J. (1978). Modeling by shortest data description. *Automatica*, 14(5):465–471.

Rizzo, P., Veloso, M., Miceli, M., and Cesta, A. (1997). Personality-driven social behaviors in believable agents. In *Proceedings of the AAAI Fall Symposium on Socially Intelligent Agents*, pages 109–114. AAAI.

Rizzolatti, G. and Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27:169–192.

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22(1):27–57.

Rohrer, B. (2010). Accelerating progress in artificial general intelligence: Choosing a benchmark for natural world interaction. *Journal of Artificial General Intelligence*, 2(1):1–28.

Roid, G. (2003). *Stanford-Binet intelligence scales*. Riverside.

Romanés, G. J. (1885). Mind in men and animals. *The North American Review*, pages 251–257.

Rönnberg, J. and Borg, E. (2001). A review and evaluation of research on the deaf-blind from perceptual, communicative, social and rehabilitative perspectives. *Scandinavian Audiology*, 30(2):67–77.

Roth, G. and Dicke, U. (2013). Evolution of nervous systems and brains. In Galizia, C. and Lledo, P., editors, *Neurosciences – from molecule to behavior: A university textbook*, pages 19–45. Springer.

Rowe, C. and Healy, S. D. (2014). Measuring cognition will be difficult but worth it: A response to comments on Rowe and Healy. *Behavioral Ecology*, 24:1298.

Rumbaugh, D. M. (2014). *Language learning by a chimpanzee: The Lana project*. Academic Press.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1985). Learning internal representations by error propagation. Technical report, DTIC Document.

Rummery, G. and Niranjan, M. (1994). *On-line Q-learning using connectionist systems*. Cambridge University Engineering Department, TR 166.

Rundquist, E. A. (1936). Intelligence test scores and school marks in 1928 and 1933. *School and Society*, 43:301–304.

Russell, S. (2015). The long-term future of (artificial) intelligence. AAAI 2015 invited talk.

Russell, S., Dewey, D., Tegmark, M., Aguirre, A., Brynjolfsson, E., Calo, R., Dietterich, T., George, D., Hibbard, B., Hassabis, D., et al. (2015). Research priorities for robust and beneficial artificial intelligence: an open letter. Future of Life Institute (FLI). http://arxiv.org/abs/1602.03506.

Russell, S. and Norvig, P. (2009). *Artificial intelligence: A modern approach*, 3rd edition. Prentice Hall.

Rust, J. and Golombok, S. (2009). *Modern psychometrics: The science of psychological assessment*. Routledge.

Ryabko, B. and Reznikova, Z. (2009). The use of ideas of information theory for studying "language" and intelligence in ants. *Entropy*, 11(4):836–853.

Rzeszotarski, J. M. and Kittur, A. (2011). Instrumenting the crowd: Using implicit behavioral measures to predict task performance. In *Proceedings of the 24th annual ACM Symposium on User Interface Software and Technology*, pages 13–22. ACM.

Saklofske, D. H. and Zeidner, M. (1995). *International handbook of personality and intelligence*. Springer Science and Business Media.

Salminen, J. (2012). Collective intelligence in humans: A literature review. CoRR arXiv preprint:1204.3401.

Salovey, P. and Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality*, 9(3):185–211.

Sanberg, P. (1976). "Neural capacity" in Mimosa pudica: A review. *Behavioral Biology*, 17(4):435–452.

Sanderson, A. C. (2000). Minimal representation size metrics for intelligent robotic systems. In Meystel, A. M. and Messina, E. R., editors, *Measuring the performance and intelligence of systems: Proceedings of the 2000 PerMIS Workshop*, pages 127–134. NIST Special Publication 970. NIST.

Sanghi, P. and Dowe, D. L. (2003). A computer program capable of passing IQ tests. In *4th International Conference on Cognitive Science (ICCS'03), Sydney*, pages 570–575. International Association for Cognitive Science.

Savage-Rumbaugh, E. S. and Lewin, R. (1994). *Kanzi: The ape at the brink of the human mind*. John Wiley.

Saygin, A. P., Cicekli, I., and Akman, V. (2000). Turing test: 50 years later. *Minds and Machines*, 10(4):463–518.

Scalise, K. and Allen, D. D. (2015). Use of open-source software for adaptive measurement: Concerto as an R-based computer adaptive development and delivery platform. *British Journal of Mathematical and Statistical Psychology*, 68(3):478–496, available at http://concertoplatform.com/.

Schaeffer, J., Burch, N., Bjornsson, Y., Kishimoto, A., Muller, M., Lake, R., Lu, P., and Sutphen, S. (2007). Checkers is solved. *Science*, 317(5844):1518.

Schaie, K. W. (2010). Primary mental abilities. In Weiner, I. B. and Craighead, W. E., editors, *Corsini encyclopedia of psychology*, pages 1286–1288. John Wiley.

Schaie, K. W., Willis, S. L., and Caskie, G. I. (2004). The Seattle longitudinal study: Relationship between personality and cognition. *Aging Neuropsychology and Cognition*, 11(2–3):304–324.

Schaul, T. (2014). An extensible description language for video games. *IEEE Transactions on Computational Intelligence and AI in Games*, 6(4):325–331.

Schenck, C. (2013). Intelligence tests for robots: Solving perceptual reasoning tasks with a humanoid robot. Master's thesis, Iowa State University.

Schenck, C., Sinapov, J., and Stoytchev, A. (2012). Which object comes next? Grounded order completion by a humanoid robot. *Cybernetics and Information Technologies*, 12(3):5–16.

Schenck, C. and Stoytchev, A. (2012). The object pairing and matching task: Toward Montessori tests for robots. In Ugur, E., Nagai, Y., Oztop, E., and Asada, M., editors, *Proceedings of Humanoids 2012 Workshop on Developmental Robotics: Can developmental robotics yield human-like cognitive abilities?*, pages 7–13. IEEE.

Scheuneman, J. (1979). A method of assessing bias in test items. *Journal of Educational Measurement*, 16(3):143–152.

Schlenoff, C., Scott, H., and Balakirsky, S. (2011). Performance evaluation of intelligent systems at the NIST. Technical report, DTIC Document.

Schmidhuber, J. (2004). Optimal ordered problem solver. *Machine Learning*, 54(3):211–254.

Schmidhuber, J. (2007). Gödel machines: Fully self-referential optimal universal self-improvers. In B. Goertzel and C. Pennachin, editors *Artificial general intelligence*, pages 199–226. Springer.

Schmitt, V., Pankau, B., and Fischer, J. (2012). Old world monkeys compare to apes in the primate cognition test battery. *PloS ONE*, 7(4):e32024.

Schönhage, D. D. A. and Strassen, V. (1971). Schnelle multiplikation grosser zahlen. *Computing*, 7(3–4):281–292.

Schrepp, M. (1999). An empirical test of a process model for letter series completion problems. In Albert, D. and Lukas, J., editors, *Knowledge spaces: Theories, empirical research, applications*, pages 133–154. Lawrence Erlbaum Associates.

Schweizer, P. (1998). The truly total Turing test. *Minds and Machines*, 8(2):263–272.

Scriven, M. (1965). An essential unpredictability in human behavior. In Wolman, B. B. and Nagel, E., editors, *Scientific psychology: Principles and approaches*, pages 411–425. Basic Books.

Sculley, D. and Brodley, C. E. (2006). Compression and machine learning: A new perspective on feature space vectors. In *Data Compression Conference*, pages 332–341. IEEE.

Seager, S. (2013). Exoplanet habitability. *Science*, 340(6132):577–581.

Sebeok, T. A. and Rosenthal, R. E. (1981). *The Clever Hans phenomenon: Communication with horses, whales, apes, and people.* New York Academy of Sciences.

Seferta, A., Guay, P.-J., Marzinotto, E., and Lefebvre, L. (2001). Learning differences between feral pigeons and zenaida doves: The role of neophobia and human proximity. *Ethology*, 107(4):281–293.

Segerdahl, P. (2012). Human-enculturated apes: Towards a new synthesis of philosophy and comparative psychology. In Birke, L. and Hockenhull, J., editors, *Crossing boundaries: Investigating human-animal relationships*, chapter 7, pages 139–159. BRILL.

Semel, E. M., Wiig, E. H., and Secord, W. (1996). *CELF 3, Clinical Evaluation of Language Fundamentals: Observational rating scales*. Psychological Corporation.

Shah, H. and Warwick, K. (2015). Human or machine? *Communications of the ACM*, 58(4):8.

Shahaf, D. and Amir, E. (2007). Towards a theory of AI completeness. In *AAAI Spring Symposium: Logical Formalizations of Commonsense Reasoning*, pages 150–155. Commonsense Reasoning.

Shahaf, D. and Horvitz, E. (2010). Generalized task markets for human and machine computation. In *National Conference on Artificial Intelligence*, pages 986–993. AAAI.

Shanahan, M. (2015). *The technological singularity*. MIT Press.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 623–656.

Shannon, C. E. (1987). Interview. *Omni Magazine*, 9(11 August):33–37.

Shapiro, S. C. (1992). *Encyclopedia of artificial intelligence*, 2nd edition. John Wiley.

Sharma, A., Vo, N. H., Gaur, S., and Baral, C. (2015). An approach to solve Winograd schema challenge using automatically extracted commonsense knowledge. In *2015 AAAI Spring Symposium Series*, pages 141–144. AAAI.

Shettleworth, S. J. (2009). The evolution of comparative cognition: Is the snark still a boojum? *Behavioural Processes*, 80(3):210–217.

Shettleworth, S. J. (2010a). Clever animals and killjoy explanations in comparative psychology. *Trends in Cognitive Sciences*, 14(11):477–481.

Shettleworth, S. J. (2010b). *Cognition, evolution, and behavior*, 2nd edition. Oxford University Press.

Shettleworth, S. J. (2013). *Fundamentals of comparative cognition*. Oxford University Press.

Shieber, S. M. (2004). *The Turing test: Verbal behavior as the hallmark of intelligence*. MIT Press.

Shipley, W. C. (1940). A self-administering scale for measuring intellectual impairment and deterioration. *Journal of Psychology*, 9(2):371–377.

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, 74(1):107–120.

Sijtsma, K. (2012). Future of psychometrics: Ask what psychometrics can do for psychology. *Psychometrika*, 77(1):4–20.

Simmons, R. (2000). Survivability and competence as measures of intelligent systems. In Meystel, A. M. and Messina, E. R., editors, *Measuring the performance and intelligence of systems: Proceedings of the 2000 PerMIS Workshop*, pages 162–163. NIST Special Publication 970. NIST.

Simon, H. A. (1978). On the forms of mental representation. *Perception and Cognition: Issues in the Foundations of Psychology*, 9:3–18.

Simon, H. A. (1980). Cognitive science: The newest science of the artificial. *Cognitive Science*, 4(1):33–46.

Simon, H. A. (1982). *Models of bounded rationality: Empirically grounded economic reason*, volume 3. MIT Press.

Simon, H. A. and Kotovsky, K. (1963). Human acquisition of concepts for sequential patterns. *Psychological Review*, 70(6):534.

Simon, T., Klahr, D., and Newell, A. (1991). SCSoar: Pattern induction in series completion problem solving. In *European Soar Workshop*, page 17. Cambridge University Press.

Simpson, R. and Twardy, C. (2008). Refining the cognitive decathlon. In Madhavan, R. and Messina, E., editors, *Proceedings of the 8th Workshop on Performance Metrics for Intelligent Systems*, pages 124–131. NIST.

Sinapov, J. and Stoytchev, A. (2010). The odd one out task: Toward an intelligence test for robots. In *2010 IEEE 9th International Conference on Development and Learning (ICDL)*, pages 126–131. IEEE.

Singer, P. (1995). *Animal liberation*. Random House.

Sirois, S. and Shultz, T. R. (2006). Preschoolers out of adults: Discriminative learning with a cognitive load. *Quarterly Journal of Experimental Psychology*, 59(8):1357–1377.

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis.* Appleton-Century.

Sloman, A. (1978). *The computer revolution in philosophy: Philosophy, science and models of mind*, volume 3. Harvester Press Brighton. Freely available version with updates, 2009, http://www.cs.bham.ac.uk/research/projects/cogaff/crp/crp.pdf.

Sloman, A. (1984). The structure and space of possible minds. In Torrance, S., editor, *The mind and the machine: Philosophical aspects of artificial intelligence*, pages 35–42. Ellis Horwood.

Sloman, A. (2006). Why Asimov's three laws of robotics are unethical. http://www.cs.bham.ac.uk/research/projects/cogaff/misc/asimov-three-laws.html.

Sloman, A. (2009). Some requirements for human-like robots: Why the recent over-emphasis on embodiment has held up progress. In *Creating brain-like intelligence*, pages 248–277. Springer.

Sloman, A. (2014a). Judging chatbots at Turing test. http://www.cs.bham.ac.uk/research/projects/cogaff/misc/turing-test-2014.html.

Sloman, A. (2014b). Judging chatbots without opening them: Limitations of 'black-box' tests. http://www.cs.bham.ac.uk/research/projects/cogaff/misc/black-box-tests.html.

Sloman, A. (2016). Construction kits for biological evolution. In Soskova, M. and Coope, S. B., editors, *The incomputable*. Springer. To appear.

Slotnick, B. (2001). Animal cognition and the rat olfactory system. *Trends in Cognitive Sciences*, 5(5):216–222.

Smith, C. (2010). *What is a person? Rethinking humanity, social life, and the moral good from the person up*. University of Chicago Press.

Smith, J. M. and Harper, D. (2003). *Animal signals*. Oxford University Press.

Smith, K. C. (2009). The trouble with intrinsic value: An ethical primer for astrobiology. In Bertka, C. M., editor, *Exploring the origin, extent, and future of life*, pages 261–280. Cambridge University Press.

Smith, K. C. (2014). Manifest complexity: A foundational ethic for astrobiology? *Space Policy*, 30(4):209–214.

Smith, W. D. (2002). Rating systems for gameplayers, and learning. Technical Report, NEC.

Smith, W. D. (2006). Mathematical definition of "intelligence" (and consequences). Unpublished report.

Snow, R. E., Kyllonen, P. C., and Marshalek, B. (1984). The topography of ability and learning correlations. *Advances in the Psychology of Human Intelligence*, 2:47–103.

Soares, C. (2009). UCI++: Improved support for algorithm selection using datasetoids. In *Advances in knowledge discovery and data mining*, pages 499–506. Springer.

Soares, N., Fallenstein, B., Armstrong, S., and Yudkowsky, E. (2015). Corrigibility. In *AAAI-15 Workshop on AI and Ethics*, pages 74–82. AAAI.

Sobel, D. (2005). *Longitude: The true story of a lone genius who solved the greatest scientific problem of his time*. Macmillan.

Sohn, K., Krishnamoorthy, S., Paul, O., and Lewis, M. A. (2012). Giving robots a flexible persona: The five factor model of artificial personality in action. In *12th International Conference on Control, Automation and Systems (ICCAS)*, pages 133–139. IEEE.

Soler-Toscano, F., Zenil, H., Delahaye, J.-P., and Gauvrit, N. (2014). Calculating Kolmogorov complexity from the output frequency distributions of small Turing machines. *PloS ONE*, 9(5):e96223.

Solomonoff, G. (2010). Ray Solomonoff (1926–2009). *Algorithms*, 3(3):255–259.

Solomonoff, R. J. (1960). A preliminary report on a general theory of inductive inference. Report V-131, Zator Co.

Solomonoff, R. J. (1962). Training sequences for mechanized induction. In Yovit, M., Jacobi, G., and Goldsteins, G., editors, *Self-organizing systems*, pages 425–434. Spartan Books.

Solomonoff, R. J. (1964a). A formal theory of inductive inference. Part I. *Information and Control*, 7(1):1–22.

Solomonoff, R. J. (1964b). A formal theory of inductive inference. Part II. *Information and Control*, 7(2):224–254.

Solomonoff, R. J. (1978). Complexity-based induction systems: Comparisons and convergence theorems. *IEEE Transactions on Information Theory*, 24(4):422– 432.

Solomonoff, R. J. (1984a). Optimum sequential search. Oxbridge Research. http://raysolomonoff.com/publications/optseq.pdf.

Solomonoff, R. J. (1984b). Perfect training sequences and the costs of corruption – a progress report on induction inference research. Oxbridge Research. http://raysolomonoff.com/publications/perfecttraining.pdf.

Solomonoff, R. J. (1985). The time scale of artificial intelligence: Reflections on social effects. *Human Systems Management*, 5:149–153.

Solomonoff, R. J. (1989). A system for incremental learning based on algorithmic probability. In *Proceedings of the Sixth Israeli Conference on Artificial Intelligence, Computer Vision and Pattern Recognition*, pages 515–527. Information Processing Association of Israel.

Solomonoff, R. J. (1996). Does algorithmic probability solve the problem of induction? In Dowe, D., Korb, K., and Oliver, J., editors, *Proceedings of the Information, Statistics and Induction in Science (ISIS) Conference, Melbourne, Australia*, pages 7–8. World Scientific.

Solomonoff, R. J. (1997). The discovery of algorithmic probability. *Journal of Computer and System Sciences*, 55(1):73–88.

Solomonoff, R. J. (2002). Progress in incremental machine learning. In *NIPS Workshop on Universal Learning Algorithms and Optimal Search, Whistler, BC*.

Soloveichik, D. and Winfree, E. (2007). Complexity of self-assembled shapes. *SIAM Journal on Computing*, 36(6):1544–1569.

Sparrow, B., Liu, J., and Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science*, 333(6043):776–778.

Sparrow, R. (2004). The Turing triage test. *Ethics and Information Technology*, 6(4):203–213.

Sparrow, R. (2012). Can machines be people? reflections on the Turing triage test. In Lin, P., Abney, K., and Bekey, G., editors, *Robot ethics: The ethical and social implications of robotics*, pages 301–315. MIT Press.

Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of Psychology*, 15(2):201–292.

Spearman, C. (1925). Some issues in the theory of "g" (including the law of diminishing returns). *Nature*, 116:436–439.

Spearman, C. (1927). *The abilities of man: Their nature and measurement*. Macmillan.

Spitz, H. H. (1986). *The raising of intelligence: A selected history of attempts to raise retarded intelligence*. Routledge.

Spoto, A., Stefanutti, L., and Vidotto, G. (2010). Knowledge space theory, formal concept analysis, and computerized psychological assessment. *Behavior Research Methods*, 42(1):342–350.

Spronck, P., Ponsen, M., Sprinkhuizen-Kuyper, I., and Postma, E. (2006). Adaptive game AI with dynamic scripting. *Machine Learning*, 63(3):217–248.

Spronck, P., Sprinkhuizen-Kuyper, I., and Postma, E. (2003). Improving opponent intelligence through offline evolutionary learning. *International Journal of Intelligent Games and Simulation*, 2(1):20–27.

Spronck, P., Sprinkhuizen-Kuyper, I., and Postma, E. (2004). Difficulty scaling of game AI. In *Proceedings of the 5th International Conference on Intelligent Games and Simulation (GAME-ON 2004)*, pages 33–37. The European Multidisciplinary Society for Modelling and Simulation Technology.

Srinivasan, R. (2002). *Importance sampling: Applications in communications and detection*. Springer.

Staiano, J., Lepri, B., Aharony, N., Pianesi, F., Sebe, N., and Pentland, A. (2012). Friends don't lie: Inferring personality traits from social network structure. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 321–330. ACM.

Stankov, L. (2000). Complexity, metacognition, and fluid intelligence. *Intelligence*, 28(2):121–143.

Starkie, B., van Zaanen, M., and Estival, D. (2006). The Tenjinno machine translation competition. In *Grammatical inference: Algorithms and applications*, pages 214–226. Springer.

Steels, L. (2011). *Design patterns in fluid construction grammar*, volume 11. John Benjamins.

Steels, L., Spranger, M., Van Trijp, R., Höfer, S., and Hild, M. (2012). Emergent action language on real robots. In *Language grounding in robots*, pages 255–276. Springer.

Steiner, I. D. (1966). Models for inferring relationships between group size and potential group productivity. *Behavioral Science*, 11(4):273–283.

Stenhouse, D. (1974). *The evolution of intelligence*. Barnes and Noble.

Stern, W. (1912). *Die psychologischen Methoden der Intelligenzprüfung und deren Anwendung an Schulkindern*. Barth.

Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities.* Lawrence Erlbaum Associates.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. Cambridge University Press.

Sternberg, R. J. (1997). A triarchic view of giftedness: Theory and practice. *Handbook of Gifted Education*, 2:43–53.

Sternberg, R. J. (1999). Intelligence as developing expertise. *Contemporary Educational Psychology*, 24(4):359–375.

Sternberg, R. J., editor (2000). *Handbook of intelligence*. Cambridge University Press.

Sternberg, R. J., Conway, B. E., Ketron, J. L., and Bernstein, M. (1981). People's conceptions of intelligence. *Journal of Personality and Social Psychology*, 41(1):37.

Sternberg, R. J. and Gardner, M. K. (1983). Unities in inductive reasoning. *Journal of Experimental Psychology: General*, 112(1):80.

Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, 103(2684):677–689.

Stewart, L., MacLean, E. L., Ivy, D., Woods, V., Cohen, E., Rodriguez, K., McIntyre, M., Mukherjee, S., Call, J., Kaminski, J., Miklósi, Á., Wrangham, R. W., and Hare, B. (2015). Citizen science as a new tool in dog cognition research. *PloS ONE*, 10(9):e0135176.

Steyn, M. (2006). It's the demography, stupid. *New Criterion*, 24(5):10.

Stone, P., Kaminka, G. A., Kraus, S., Rosenschein, J. S. (2010). Ad hoc autonomous agent teams: Collaboration without pre-coordination. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence*, pages 1504–1509. AAAI.

Stork, D. G. (1998). Scientist on the set: An interview with Marvin Minsky. In *HAL's legacy: 2001's computer as dream and reality*, pages 15–32. MIT Press.

Strassmann, J. E., Klingler, C. J., Arévalo, E., Zacchi, F., Husain, A., Williams, J., Seppä, P., and Queller, D. C. (1997). Absence of within-colony kin discrimination in behavioural interactions of swarm-founding wasps. *Proceedings of the Royal Society of London, Series B*, 264(1388):1565.

Sturtevant, N. (2012). Benchmarks for grid-based pathfinding. *Transactions on Computational Intelligence and AI in Games*, 4(2):144–148.

Sutcliffe, G. (2009). The TPTP Problem Library and associated infrastructure: The FOF and CNF Parts, v3.5.0. *Journal of Automated Reasoning*, 43(4):337–362.

Sutcliffe, G. and Suttner, C. (2006). The state of CASC. *AI Communications*, 19(1):35–48.

Suthana, N. and Fried, I. (2014). Deep brain stimulation for enhancement of learning and memory. *Neuroimage*, 85:996–1002.

Sutton, R. S. and Barto, A. G. (1998). *Reinforcement learning: An introduction*. MIT Press.

Svartberg, K. and Forkman, B. (2002). Personality traits in the domestic dog (*Canis familiaris*). *Applied Animal Behaviour Science*, 79(2):133–155.

Swain, D. T., Couzin, I. D., and Leonard, N. E. (2012). Real-time feedback-controlled robotic fish for behavioral experiments with fish schools. *Proceedings of the IEEE*, 100(1):150–163.

Tannenbaum, A. J. (1986). Giftedness: A psychosocial approach. In Sternberg, R. and Davidson, J., editors, *Conceptions of giftedness*, pages 21–52. Cambridge University Press.

Tarapore, D., Lungarella, M., and Gómez, G. (2006). Quantifying patterns of agent–environment interaction. *Robotics and Autonomous Systems*, 54(2):150– 158.

Teng, S.-Y. (1943). Chinese influence on the Western examination system: I. Introduction. *Harvard Journal of Asiatic Studies*, 7(4):267–312.

Terman, L. M. (1922). The great conspiracy or the impulse imperious of intelligence testers, psychoanalyzed and exposed by Mr. Lippmann. *New Republic*, 33(27 December):116–120.

Terrace, H. S. (1987). *Nim: A chimpanzee who learned sign language*. Columbia University Press.

Tesauro, G. (1994). TD-Gammon, a self-teaching backgammon program, achieves master-level play. *Neural Computation*, 6(2):215–219.

Thagard, P. R. (1978). The best explanation: Criteria for theory choice. *Journal of Philosophy*, 75(2):76–92.

Theiner, G., Allen, C., and Goldstone, R. L. (2010). Recognizing group cognition. *Cognitive Systems Research*, 11(4):378–395.

Thomas, R. K. (1996). Investigating cognitive abilities in animals: Unrealized potential. *Cognitive Brain Research*, 3(3):157–166.

Thompson, J. (2013). Special issue: The Flynn effect re-evaluated. *Intelligence*, 41(6):751–857.

Thompson, P. M., Cannon, T. D., Narr, K. L., Van Erp, T., Poutanen, V.-P., Huttunen, M., Lönnqvist, J., Standertskjöld-Nordenstam, C.-G., Kaprio, J., Khaledy, M., et al. (2001). Genetic influences on brain structure. *Nature Neuroscience*, 4(12):1253–1258.

Thompson, S. K. (2012). *Sampling*, 3rd edition. John Wiley.

Thórisson, K. R., Bieger, J., Schiffel, S., and Garrett, D. (2015). Towards flexible task environments for comprehensive evaluation of artificial intelligent systems and automatic learners. In Bieger, J., Goertzel, B., and Potapov, A., editors, *International Conference on Artificial General Intelligence*, pages 187–196. Springer.

Thorndike, E. L. (1911). *Animal intelligence: Experimental studies*. Macmillan.

Thorndike, E. L. (1920). Intelligence and its uses. *Harper's Magazine*, 140:227–335.

Thorndike, E. L., Bregman, E. O., Cobb, M. V., and Woodyard, E. (1927). *The measurement of intelligence*. Teachers College Bureau of Publications, Columbia University.

Thorndike, R. L. (1935). *Organization of behavior in the albino rat*. Genetic Psychology Monographs. Heldref.

Thorndike, R. L. (1936). Factor analysis of social and abstract intelligence. *Journal of Educational Psychology*, 27(3):231.

Thornton, A., Isden, J., and Madden, J. R. (2014). Toward wild psychometrics: Linking individual cognitive differences to fitness. *Behavioral Ecology*, aru095.

Thornton, A. and Lukas, D. (2012). Individual variation in cognitive performance: Developmental and evolutionary perspectives. *Philosophical Transactions of the Royal Society, Series B*, 367(1603):2773–2783.

Thorp, T. R. and Mahrer, A. R. (1959). Predicting potential intelligence. *Journal of Clinical Psychology*, 15(3):286–288.

Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology*, 33:529–554.

Thurstone, L. L. (1929). Theory of attitude measurement. *Psychological Review*, 36(3):222–241.

Thurstone, L. L. (1937). Ability, motivation, and speed. *Psychometrika*, 2(4):249– 254.

Thurstone, L. L. (1938). Primary mental abilities. *Psychometric Monographs*.

Thurstone, L. L. and Thurstone, T. G. (1941). *Factorial studies of intelligence*. Psychometrika Monograph Supplements. University of Chicago Press.

Thurstone, L. L. and Thurstone, T. G. (1947). *American Council on Education for college freshmen: Manual of instructions*. Cooperative Test Division, Educational Testing Service.

Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie*, 20(4):410–433.

Togelius, J., Yannakakis, G. N., Karakovskiy, S., and Shaker, N. (2012). Assessing believability. In Hingston, P., editor, *Believable bots*, pages 215–230. Springer.

Tomasello, M. (2006). Acquiring linguistic constructions, volume two. In Damon, W. and Lerner, R. M., editors, *Handbook of child psychology*, 255–298. John Wiley.

Tomasello, M. and Call, J. (1997). *Primate cognition*. Oxford University Press.

Tomasello, M. and Carpenter, M. (2007). Shared intentionality. *Developmental Science*, 10(1):121–125.

Tomasik, B. (2014). Do artificial reinforcement-learning agents matter morally? CoRR arXiv preprint:1410.8233.

Torres-Quevedo, L. (1915a). Ensayos sobre automática – su definición. Extensión teórica de sus aplicaciones. *Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales 12, 391–418. English translation, Randell, B., editor. The origins of digital computers. Springer*, 1973.

Torres-Quevedo, L. (1915b). Torres and his remarkable automatic devices. *Scientific American*, 80(2079):296–298.

Treffert, D. A. (2009). The savant syndrome: An extraordinary condition. A synopsis: Past, present, future. *Philosophical Transactions of the Royal Society, Series B*, 364(1522):1351–1357.

Trewavas, A. (2005). Plant intelligence. *Naturwissenschaften*, 92(9):401–413.

Tromp, J. (2006). Binary lambda calculus and combinatory logic. In M. Hutter, W. Merkle and P. M. B. Vitányi *Kolmogorov Complexity and Applications*, Dagstuhl Seminar Proceedings 6051.

Tucker-Drob, E. M. (2009). Differentiation of cognitive abilities across the life span. *Developmental Psychology*, 45(4):1097.

Tummolini, L., Castelfranchi, C., Ricci, A., Viroli, M., and Omicini, A. (2005). "Exhibitionists" and "Voyeurs" do it better: A shared environment for flexible coordination with tacit messages. In *Environments for multi-agent systems*, pages 215–231. Springer.

Turing, A. M. (1936). On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society*, 2(42):230–265.

Turing, A. M. (1937). On computable numbers, with an application to the entscheidungsproblem: A correction. *Proceedings of the London Mathematical Society*, 2(6):544–456.

Turing, A. M. (1948). Intelligent machinery. Chapter 10 in B. J. Copeland, editor, *The essential Turing: Seminal writings in computing, logic, philosophy, artificial intelligence, and artificial life: Plus the secrets of enigma*. Reprint, Oxford University Press, 2004.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59:433–460.

Turing, A. M. (1951). Intelligent machinery, a heretical theory. Chapter 12 in B. J. Copeland, editor, *The essential Turing: Seminal writings in computing, logic, philosophy, artificial intelligence, and artificial life: Plus the secrets of enigma*. Reprint, Oxford University Press, 2004.

Turing, A. M. (1952). Can automatic calculating machines be said to think? BBC Third Programme, 14 and 23 Jan. 1952, between M.H.A. Newman, AMT, Sir Geoffrey Jefferson and R.B. Braithwaite, in Copeland, B. J., editor, *The essential Turing*, pp. 494–495. Oxford University Press. http://www.turingarchive.org/browse.php/B/6.

Turing, A. M. and Champernowne, D. (1948). Turochamp. In Copeland, B. J., editor, *The essential Turing*, pp. 562–575. Oxford University Press.

Turney, P. D. (2011). Analogy perception applied to seven tests of word comprehension. *Journal of Experimental and Theoretical Artificial Intelligence*, 23(3):343–362.

Tuyls, K. and Weiss, G. (2012). Multiagent learning: Basics, challenges, and prospects. *AI Magazine*, 33(3):41.

Tyler, T. (2011). The matching pennies tournament. http://matchingpennies.com/tournament/.

Úbeda, Y. and Llorente, M. (2014). Personality in sanctuary-housed chimpanzees: A comparative approach of psychobiological and penta-factorial human models. *Evolutionary Psychology*, 13(1):182–196.

Ulam, S. (1958). Tribute to John von Neumann. *Bulletin of the American Mathematical Society*, 64(3):1–49.

Ulinwa, V. C. I. (2008). *Machine Intelligence Quotient*. VDM.

Urbina, S. (2011). Tests of intelligence. In Sternberg, R. J. and Kaufman, S. B., editors, *The Cambridge handbook of intelligence*, pages 20–38. Cambridge University Press.

Vakoch, D. A. (2011). *Communication with extraterrestrial intelligence*. SUNY Press.

Valiant, L. G. (1984). A theory of the learnable. *Communications of the ACM*, 27(11):1134–1142.

Valiant, L. G. (2013). *Probably approximately correct: Nature's algorithms for learning and prospering in a complex world*. Basic Books.

Van Der Linden, W. J. (2008). Using response times for item selection in adaptive testing. *Journal of Educational and Behavioral Statistics*, 33(1):5–20.

Van Der Maas, H. L. J., Dolan, C. V., Grasman, R. P., Wicherts, J. M., Huizenga, H. M., and Raijmakers, M. E. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review*, 113(4): 842.

Van Der Maas, H. L. J., Kan, K.-J., and Borsboom, D. (2014). Intelligence is what the intelligence test measures. Seriously. *Journal of Intelligence*, 2(1):12–15.

Van Duijn, M., Keijzer, F., and Franken, D. (2006). Principles of minimal cognition: Casting cognition as sensorimotor coordination. *Adaptive Behavior*, 14(2):157–170.

Van Heijningen, C. A., De Visser, J., Zuidema, W., and Ten Cate, C. (2009). Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. *Proceedings of the National Academy of Sciences*, 106(48):20538–20543.

Van Knippenberg, D., De Dreu, C. K., and Homan, A. C. (2004). Work group diversity and group performance: An integrative model and research agenda. *Journal of Applied Psychology*, 89(6):1008.

Van Knippenberg, D. and Schippers, M. C. (2007). Work group diversity. *Annual Review of Psychology*, 58:515–541.

Van Rijn, J. N., Bischl, B., Torgo, L., Gao, B., Umaashankar, V., Fischer, S., Winter, P., Wiswedel, B., Berthold, M. R., and Vanschoren, J. (2013). OpenML: A collaborative science platform. In *European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 645–649. *Lecture Notes in Artificial Intelligence*, volume 8190. Springer.

Vanschoren, J., Blockeel, H., Pfahringer, B., and Holmes, G. (2012). Experiment databases. *Machine Learning*, 87(2):127–158.

Vanschoren, J., Van Rijn, J. N., Bischl, B., and Torgo, L. (2014). Openml: Networked science in machine learning. *ACM SIGKDD Explorations Newsletter*, 15(2):49–60.

Vardi, M. Y. (2012). Artificial intelligence: Past and future. *Communications of the ACM*, 55(1):5.

Vardi, M. Y. (2015). Human or machine? Response. *Communications of the ACM*, 58(4):8–9.

Varela, F. J., Thompson, E., and Rosch, E. (1992). *The embodied mind: Cognitive science and human experience*. MIT Press.

Vázquez, D., López, A. M., Marín, J., Ponsa, D., and Gerónimo, D. (2014). Virtual and real world adaptation for pedestrian detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(4):797–809.

Veness, J., Ng, K., Hutter, M., and Silver, D. (2011). A Monte Carlo AIXI approximation. *Journal of Artificial Intelligence Research*, 40:95–142.

Vere, S. A. (1992). A cognitive process shell. *Behavioral and Brain Sciences*, 15(03):460–461.

Vernon, M., et al. (1979). Psychological evaluation and testing of children who are deaf-blind. *School Psychology Digest*, 8(3):291–295.

Vernon, P. E. (1950). *The structure of human abilities.* Methuen.

Veselka, L., Just, C., Jang, K. L., Johnson, A. M., and Vernon, P. A. (2012). The general factor of personality: A critical test. *Personality and Individual Differences*, 52(3):261–264.

Vickrey, C. and Neuringer, A. (2000). Pigeon reaction time, Hick's law, and intelligence. *Psychonomic Bulletin and Review*, 7(2):284–291.

Vinciarelli, A. and Mohammadi, G. (2014). A survey of personality computing. *IEEE Transactions on Affective Computing*, 5(3):273–291.

Vinge, V. (1993). The coming technological singularity: How to survive in the post-human era. In *Vision 21: Interdisciplinary science and engineering in the era of cyberspace*, volume 1, pages 11–22. NASA.

Vitányi, P. and Li, M. (1997). On prediction by data compression. In *Machine learning: ECML-97*, pages 14–30. Springer.

Vitányi, P. M., Balbach, F. J., Cilibrasi, R. L., and Li, M. (2009). Normalized information distance. In *Information theory and statistical learning*, pages 45–82. Springer.

Von Ahn, L. (2005). *Human computation*. PhD thesis, Carnegie Mellon University.

Von Ahn, L. (2009). Human computation. In *Design Automation Conference, 2009. DAC'09. 46th ACM/IEEE*, pages 418–419. IEEE.

von Ahn, L., Blum, M., and Langford, J. (2004). Telling humans and computers apart automatically. *Communications of the ACM*, 47(2):56–60.

Von Ahn, L., Maurer, B., McMillen, C., Abraham, D., and Blum, M. (2008). RECAPTCHA: Human-based character recognition via web security measures. *Science*, 321(5895):1465.

Von Neumann, J. (1966). *Theory of self-reproducing automata*. Burks, A. W., editor. University of Illinois Press.

Vonk, J. and Beran, M. J. (2012). Bears 'count' too: Quantity estimation and comparison in black bears, *Ursus americanus*. *Animal Behaviour*, 84(1):231–238.

Voss, P. (2007). Essentials of general intelligence: The direct path to artificial general intelligence. In B. Goertzel and C. Pennachin, editors *Artificial general intelligence*, pages 131–157. Springer.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wainer, H. (2000). *Computerized adaptive testing: A primer*, 2nd edition. Lawrence Erlbaum Associates.

Wallace, A. R. (1870). The limits of natural selection as applied to man. In *Contributions to the theory of natural selection*, pages 332–371. Macmillan.

Wallace, A. R. (1904). *Man's place in the universe*. Chapman and Hall.

Wallace, C. S. (1964). A suggestion for a fast multiplier. *IEEE Transactions on Electronic Computers*, EC-13(1):14–17.

Wallace, C. S. (2005). *Statistical and inductive inference by minimum message length*. Springer.

Wallace, C. S. and Boulton, D. M. (1968). An information measure for classification. *Computer Journal*, 11(2):185–194.

Wallace, C. S. and Dowe, D. L. (1999). Minimum message length and Kolmogorov complexity. *Computer Journal*, 42(4):270–283.

Wang, G., Mohanlal, M., Wilson, C., Wang, X., Metzger, M., Zheng, H., and Zhao, B. Y. (2012). Social Turing tests: Crowdsourcing sybil detection. CoRR arXiv preprint:1205.3856.

Wang, G., Wang, T., Zheng, H., and Zhao, B. Y. (2014). Man vs. machine: Practical adversarial detection of malicious crowdsourcing workers. In *23rd USENIX Security Symposium, USENIX Association, CA*, pages 239–254. USENIX Association.

Wang, H., Gao, B., Bian, J., Tian, F., and Liu, T.-Y. (2015). Solving verbal comprehension questions in IQ test by knowledge-powered word embedding. CoRR arXiv preprint:1505.07909.

Wang, L. (2009a). *Behavioral biometrics for human identification: Intelligent applications*. IGI Global.

Wang, P. (2010). The evaluation of AGI systems. In Hutter, M., Baum, E., and Kitzelmann, E., editors, *3rd Conference on Artificial General Intelligence*, pages 164–169. Atlantis Press.

Wang, Y. (2009b). On abstract intelligence: Toward a unifying theory of natural, artificial, machinable, and computational intelligence. *International Journal of Software Science and Computational Intelligence*, 1(1):1–17.

Warwick, K. (2000). *QI: The quest for intelligence*. Piatkus.

Warwick, K. (2014). Turing Test success marks milestone in computing history. University or Reading Press Release, 8 June.

Wasserman, E. A. and Zentall, T. R. (2006). *Comparative cognition: Experimental explorations of animal intelligence*. Oxford University Press.

Watanabe, S. (1972). Pattern recognition as information compression. In *Frontiers of pattern recognition*, pages 561–567. Academic Press.

Waterhouse, L. (2006). Multiple intelligences, the Mozart effect, and emotional intelligence: A critical review. *Educational Psychologist*, 41(4):207–225.

Watkins, C. J. C. H. and Dayan, P. (1992). Q-learning. *Machine Learning*, 8(3):279– 292.

Watson, J. B. (1967). *Behavior: An introduction to comparative psychology*. Henry Holt.

Watt, S. (1996). Naive psychology and the inverted Turing test. *Psycoloquy*, 7(14).

Watt, S. (2008). Can people think? Or machines? A unified protocol for Turing testing. In Epstein, R., Roberts, G., and Beber, G., editors, *Parsing the Turing test: Philosophical and methodological issues in the quest for the thinking computer*, pages 301–318. Springer.

Webb, B. (2009). Animals versus animats: Or why not model the real iguana? *Adaptive Behavior*, 17(4):269–286.

Wechsler, D. (1944). *The measurement of intelligence*. Williams and Wilkins.

Wechsler, D. (2008). Wechsler Adult Intelligence Scale–fourth edition (WAIS-IV). NCS Pearson.

Wegner, D. M. and Gray, K. (2016). *The mind club: Who thinks, what feels, and why it matters*. Viking.

Weiss, D. J. (2011). Better data from better measurements using computerized adaptive testing. *Journal of Methods and Measurement in the Social Sciences*, 2(1):1–27.

Weiss, G. and Dillenbourg, P. (1999). What is 'multi' in multi-agent learning. In Dillenbourg, P., editor, *Collaborative learning: Cognitive and computational approaches*, pages 64–80. Pergamon Press.

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.

Wellman, M., Reeves, D., Lochner, K., and Vorobeychik, Y. (2004). Price prediction in a trading agent competition. *Journal of Artificial Intelligence Research*, 21:19–36.

Weng, J. (2000). Autonomous mental development and performance metrics for intelligent systems. In Meystel, A. M. and Messina, E. R., editors, *Measuring the*

performance and intelligence of systems: Proceedings of the 2000 PerMIS Workshop, pages 349–358. NIST Special Publication 970. NIST.

Weng, J. (2009). Task muddiness, intelligence metrics, and the necessity of autonomous mental development. *Minds and Machines*, 19(1):93–115.

Weng, J. (2013). *Natural and artificial intelligence: Introduction to computational brain-mind*. BMI Press.

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

Weston, J., Bordes, A., Chopra, S., and Mikolov, T. (2015). Towards AI-complete question answering: A set of prerequisite toy tasks. CoRR arXiv preprint:1502.05698.

Weyns, D. and Holvoet, T. (2005). On the role of environments in multiagent systems. *Informatica*, 29(4):409–421.

Whewell, W. (1847). *The philosophy of the inductive sciences*. Reprint, Cambridge University Press, 2014.

Whitby, B. (1996). Why the Turing test: AI's biggest blind alley. In Millican, P. and Clark, A., editors, *Machines and thought: The legacy of Alan Turing)*, pp. 53–63. Oxford University Press.

White, D. R., McDermott, J., Castelli, M., Manzoni, L., Goldman, B. W., Kronberger, G., Jaśkowski, W., O'Reilly, U.-M., and Luke, S. (2013). Better GP benchmarks: Community survey results and proposals. *Genetic Programming and Evolvable Machines*, 14:3–29.

Whiten, A. and Byrne, R. W. (1997). *Machiavellian intelligence II: Extensions and evaluations*, volume 2. Cambridge University Press.

Whiten, A., McGuigan, N., Marshall-Pescini, S., and Hopper, L. M. (2009). Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Philosophical Transactions of the Royal Society, Series B*, 364(1528):2417–2428.

Whiteson, S., Tanner, B., Taylor, M. E., and Stone, P. (2011). Protecting against evaluation overfitting in empirical reinforcement learning. In *Adaptive dynamic programming and reinforcement learning (ADPRL)*, pages 120–127. IEEE.

Whiteson, S., Tanner, B., and White, A. (2010). The reinforcement learning competitions. *AI Magazine*, 31(2):81–94.

Wiering, M. (2005). QV (λ)-learning: A new on-policy reinforcement learning algorithm. In D. Leone, editor, *7th European Workshop on Reinforcement Learning*, pages 17–18. EWRL.

Wiggins, J. S. (1973). *Personality and prediction: Principles of personality assessment*. Addison-Wesley.

Wiig, E. H. and Secord, W. (1989). *Test of Language Competence (TLC)*, expanded edition. Psychological Corporation.

Wilhelm, O. (2005). Measuring reasoning ability. In *Handbook of understanding and measuring intelligence*, pages 373–392. Sage.

Wilkins, J. (1668). *An essay towards a real character, and a philosophical language*. S. Gellibrand.

Williams, M. A. (2012). Robot social intelligence. In Ge, S., Khatib, O., Cabibihan, J.-J., Simmons, R., and Williams, M., editors, *Proceedings for the 4th International Conference Social Robotics, ICSR 2012*, pages 45–55. *Lecture Notes in Computer Science*, volume 7621. Springer.

Williams, P. L. and Beer, R. D. (2010). Information dynamics of evolved agents. In *From Animals to Animats 11*, pages 38–49. Springer.

Williams, R. L. (2013). Overview of the Flynn effect. *Intelligence*, 41(6):753–764.

Wilson, M., Hanlon, R. T., Tyack, P. L., and Madsen, P. T. (2007). Intense ultrasonic clicks from echolocating toothed whales do not elicit anti-predator responses or debilitate the squid *Loligo pealeii*. *Biology Letters*, 3(3):225–227.

Wilson, S. W. (1991). The animat path to AI. In Meyer, J.-A. and Wilson, S., editors, *From animals to animats*, pages 15–21. MIT Press.

Wimmer, H. and Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1):103–128.

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

Woergoetter, F. and Porr, B. (2008). Reinforcement learning. *Scholarpedia*, 3(3):1448.

Wolfram, S. (2002). *A new kind of science*. Wolfram Media.

Wolpert, D. H. (1996). The lack of a priori distinctions between learning algorithms. *Neural Computation*, 8(7):1341–1390.

Wolpert, D. H. (2012). What the no free lunch theorems really mean; how to improve search algorithms. Technical Report. Santa Fe Institute.

Wolpert, D. H. and Macready, W. G. (1995). No free lunch theorems for search. Technical Report, SFI-TR-95-02-010. Santa Fe Institute.

Wolpert, D. H. and Macready, W. G. (2005). Coevolutionary free lunches. *IEEE Transactions on Evolutionary Computation*, 9(6):721–735.

Woodley, M. A. and Bell, E. (2011). Is collective intelligence (mostly) the general factor of personality? A comment on Woolley, Chabris, Pentland, Hashmi and Malone (2010). *Intelligence*, 39(2):79–81.

Woodley, M. A., Fernandes, H. B., and Hopkins, W. D. (2015). The more g-loaded, the more heritable, evolvable, and phenotypically variable: Homology with humans in chimpanzee cognitive abilities. *Intelligence*, 50:159–163.

Woodrow, H. (1946). The ability to learn. *Psychological Review*, 53(3):147.

Woods, S., Dautenhahn, K., Kaouri, C., Boekhorst, R., and Koay, K. L. (2005). Is this robot like me? Links between human and robot personality traits. In *5th IEEE-RAS International Conference on Humanoid Robots*, pages 375–380. IEEE.

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., and Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science*, 330(6004):686–688.

WPSA (2014). Universal declaration on animal welfare (UDAW). World Society for the Protection of Animals. http://www.udaw.org/.

Wray, R. E. and Lebiere, C. (2007). Metrics for cognitive architecture evaluation. In Kaminka, G. A., editor, *Proceedings of the AAAI-07 Workshop on Evaluating Architectures for Intelligence*, pages 60–66. AAAI.

Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding, and selection in evolution. In *Proceedings of the VI International Congress of Genetics*, pages 356–366. Botanic Garden, New York.

Wright, S. (1988). Surfaces of selective value revisited. *American Naturalist*, pages 115–123.

Yampolskiy, R. V. (2012). AI-complete, AI-hard, or AI-easy–classification of problems in AI. In *The 23rd Midwest Artificial Intelligence and Cognitive Science Conference*, pages 94-101.

Yampolskiy, R. V. (2015a). *Artificial superintelligence: A futuristic approach*. CRC Press.

Yampolskiy, R. V. (2015b). The space of possible mind designs. In Bieger, J., Goertzel, B., and Potapov, A., editors, *8th International Conference on Artificial general intelligence – AGI 2015, Berlin, Germany, July 22–25, 2015, Proceedings*, pages 218–227. Springer.

Yampolskiy, R. V. and Gavrilova, M. L. (2012). Artimetrics: Biometrics for artificial entities. *Robotics and Automation Magazine*, 19(4):48–58.

Yampolskiy, R. V. and Govindaraju, V. (2008). Behavioural biometrics: A survey and classification. *International Journal of Biometrics*, 1(1):81–113.

Yockey, H. P. (2005). *Information theory, evolution, and the origin of life*. Cambridge University Press.

Yonck, R. (2012). Toward a standard metric of machine intelligence. *World Future Review*, 4(2):61–70.

York, W. and Swan, J. (2012). Taking Turing seriously (but not literally). In Muller, V. and Ayesh, A., editors, *AISB/IACAP 2012 Symposium "Revisiting Turing and His Test"*, pages 54–59. Society for the Study of Artificial Intelligence and Simulation of Behaviour.

You, J. (2015). Beyond the Turing test. *Science*, 347(6218):116.

Youyou, W., Kosinski, M., and Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences*, 112(4):1036–1040.

Yu, Y., Pan, G., Gong, Y., Xu, K., Zheng, N., Hua, W., Zheng, X., and Wu, Z. (2016). Intelligence-augmented rat cyborgs in maze solving. *PloS ONE*, 11(2):e0147754.

Yudkowsky, E. (2008). Artificial intelligence as a positive and negative factor in global risk. *Global Catastrophic Risks*, 1:303.

Zadeh, L. A. (1994). Fuzzy logic, neural networks, and soft computing. *Communications of the ACM*, 37(3):77–84.

Zadeh, L. A. (2000). The search for metrics of intelligence–a critical view. In Meystel, A. M. and Messina, E. R., editors, *Measuring the performance and intelligence of systems: Proceedings of the 2000 PerMIS Workshop*, pages 105–111. NIST Special Publication 970. NIST.

Zadeh, L. A. (2002). In quest of performance metrics for intelligent systems – A challenge that cannot be met with existing methods. In Meystel, A. M. and Messina, E. R., editors, *Measuring the performance and intelligence of systems: Proceedings of the 2000 PerMIS Workshop*, pages 303–306. NIST Special Publication 970. NIST.

Zadeh, L. A. (2008). Toward human level machine intelligence – Is it achievable? The need for a paradigm shift. *IEEE Computational Intelligence Magazine*, 3(3):11–22.

Zatuchna, Z. and Bagnall, A. (2009). Learning mazes with aliasing states: An LCS algorithm with associative perception. *Adaptive Behavior*, 17(1):28–57.

Zentall, T. R. (2012). Perspectives on observational learning in animals. *Journal of Comparative Psychology*, 126(2):114.

Zentall, T. R., Hogan, D. E., and Edwards, C. A. (1980). Oddity learning in the pigeon: Effect of negative instances, correction, and number of incorrect alternatives. *Animal Learning and Behavior*, 8(4):621–629.

Zentall, T. R., Hogan, D. E., and Holder, J. (1974). Comparison of two oddity tasks with pigeons. *Learning and Motivation*, 5(1):106–117.

Zhang, H., Horvitz, E., Miller, R. C., and Parkes, D. C. (2011). Crowdsourcing general computation. In *Proceedings of ACM CHI Conference on Human Factors*. ACM.

Zillich, M. (2012). My robot is smarter than your robot: On the need for a total Turing test for robots. In Muller, V. and Ayesh, A., editors, *AISB/IACAP 2012 Symposium*

"Revisiting Turing and His Test", pages 12–15. Society for the Study of Artificial Intelligence and Simulation of Behaviour.

Zimmermann, J., Henze, H. H., and Cremers, A. B. (2015). Gödel agents in a scalable synchronous agent framework. In Bieger, J., Goertzel, B., and Potapov, A., editors, 8th International Conference on *Artificial general intelligence*, pages 404–413. Springer.

Zitnick, C. L., Agrawal, A., Antol, S., Mitchell, M., Batra, D., and Parikh, D. (2016). Measuring machine intelligence through visual question answering. *AI* Magazine, 37(1):63–72.

Zuberbühler, K. (2000). Interspecies semantic communication in two forest primates. *Proceedings of the Royal Society of London, Series B*, 267(1444):713–718.

Zvonkin, A. K. and Levin, L. A. (1970). The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms. *Russian Mathematical Surveys*, 25:83–124.

Zwaan, R. A. and Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123(2):162.