AI and Brain: Copy or Competitor?

From "Ghost in the Machine" to Alien Intelligence: Why AI Simulates but Doesn't "Think"

Dr. Josef Sawetz, January 2026

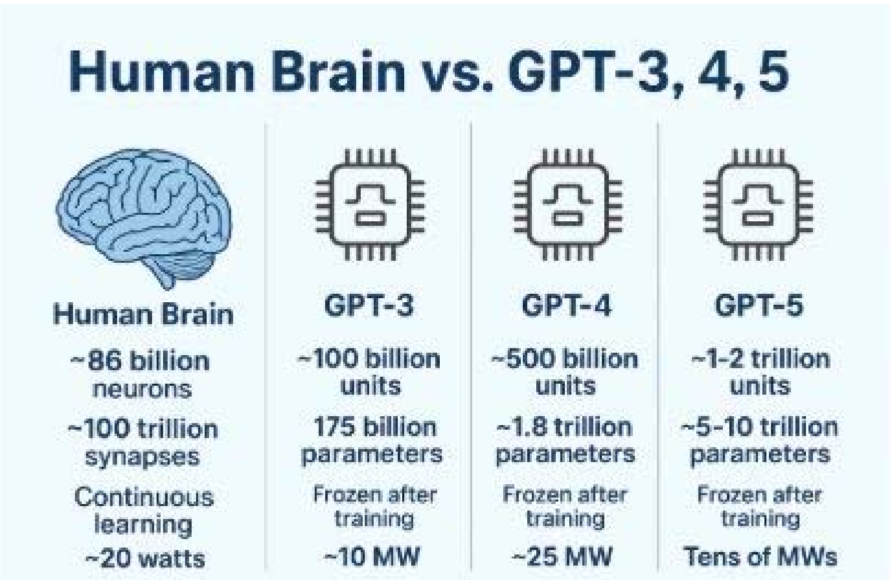
From the Illusion of Similarity to the Reality of "Alien Intelligence": A Deep Dive into the Engine Room of Thought.

It is tempting to see the ghost in the machine. When ChatGPT writes a poetic text or solves complex logic puzzles, we tend to humanize this performance. We speak of "learning," "understanding," or "feeling." However, a current analysis of cognitive architectures reveals a completely different picture: We are not dealing with a replication of the human mind, but with a "convergence of surface at divergence of depth." (Weizenbaum, 1966; Reeves & Nass, 1996; Epley et al., 2007; Gray et al., 2007; Waytz et al., 2010; Shanahan, 2022)

What does this mean? On the surface, the results look similar. In the "engine room," however, completely different mechanisms are at work. While AI is a high-performance statistical machine, the human brain remains a biological miracle of efficiency and meaning. Here is the truth about the differences that often get lost in the hype. (Hasson et al., 2020; Lake et al., 2017; Marcus, 2018)

1. The 20-Watt Miracle vs. The Power Plant: How We Really Learn

The perhaps most fundamental difference lies in the economy of thought. Your brain requires about 20 watts of energy—as much as a dim light bulb—to create world models, process feelings, and plan complex actions. Modern AI models, on the other hand, consume the electricity of entire small towns to simulate similar outputs. Yet, it is not just about electricity; it is about the method. (Attwell & Laughlin, 2001; Levy & Calvert, 2021; Strubell et al., 2019; Schwartz et al., 2020; Patterson et al., 2021)



Hebbian Plasticity vs. Backpropagation

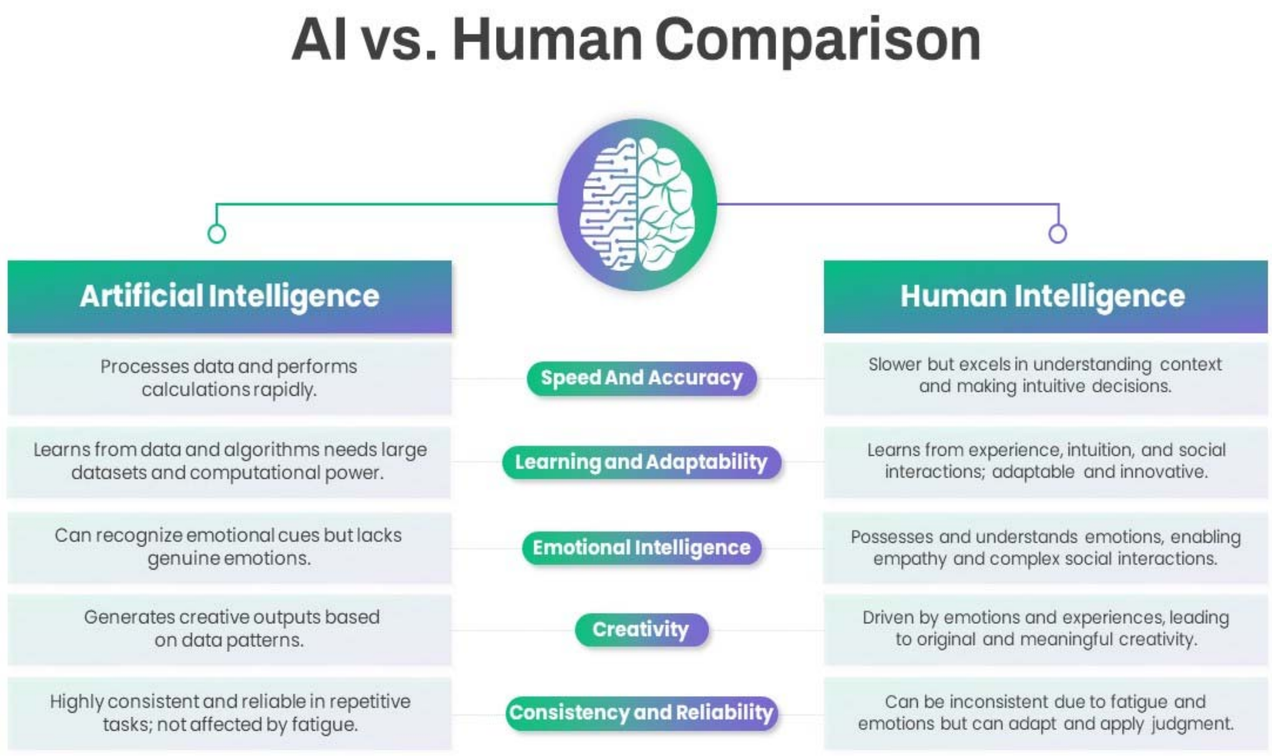
In your head, the rule applies: "Cells that fire together, wire together" (Hebbian learning). Learning happens locally, directly at the synapse, often asynchronously and extremely efficiently. If you burn your hand on a hot stove once, your brain learns immediately (One-Shot Learning). Artificial neural networks, conversely, use Backpropagation (error back-transmission). They require a global error signal that is mathematically expensively calculated back through hundreds of layers. This is biologically absolutely implausible—your brain has no "wires" to send errors backwards through the entire network. This makes AI extremely "data-hungry": It needs thousands of examples where one suffices for you. (Hebb, 1949; Bi & Poo, 1998; Lake et al., 2015; Rumelhart et al., 1986; Lillicrap et al., 2020; Marcus, 2018)

2. The Master of Fillers: Why AIs Are Only Acting

Here we arrive at a point that is often misunderstood: What is the AI actually doing when it "thinks"? Science calls this *interpolation*. Imagine you have a sheet of paper full of dots (data). The AI is a master at drawing a complex curve that connects these dots perfectly. This is called "Direct Fit." Within the known space (the training data), it is unbeatable. (Hasson et al., 2020)

The Problem of Extrapolation

But what happens when we leave the sheet of paper? Here, the machine fails. The human brain is a master of *extrapolation*. We understand the *rules* behind the dots. Example: If you know what "to walk" means and what "slowly" means, you immediately understand the concept of "walking slowly"—even if you have never seen it. We abstract principles. The AI reality: It often only memorizes statistical patterns. If forced to operate outside its trained data (Out-of-Distribution), its performance often collapses dramatically. It is, to put it provocatively, a highly intelligent copy machine that calculates probabilities but understands no causalities. (Lake et al., 2017; Marcus, 2018; Hendrycks & Dietterich, 2019; Geirhos et al., 2020; Pearl & Mackenzie, 2018; Chollet, F. 2019).



3. Remix vs. Revolution: Where the Spark is Missing

When AI paints pictures or composes music, it appears creative. But is it? We must distinguish between two types here: (Boden, 2004)

- **Combinatorial Creativity:** Re-connecting the known. Here, the AI is brilliant. It can mix the style of Van Gogh with Cyberpunk because it possesses huge databases. (Boden, 2004)
- **Transformational Creativity:** Breaking the rules to create something fundamentally new (a paradigm shift). (Boden, 2004)

True human insight—the famous "Aha!" moment—is neurologically measurable (a gamma burst in the brain) and emotionally rewarding (dopamine). It is a phase transition, a sudden restructuring of the problem. AIs do not experience "Aha!" moments. They merely optimize an error function (Loss Function) downward. They lack the *will* to break, the rebellion against the rule, and the intrinsic curiosity that drives us to explore things that have no immediate utility. (Jung-Beeman et al., 2004; Kounios & Beeman, 2009; Schultz, 1998)

4. The Simulation of Emotion: Why ChatGPT Does Not Cry With You

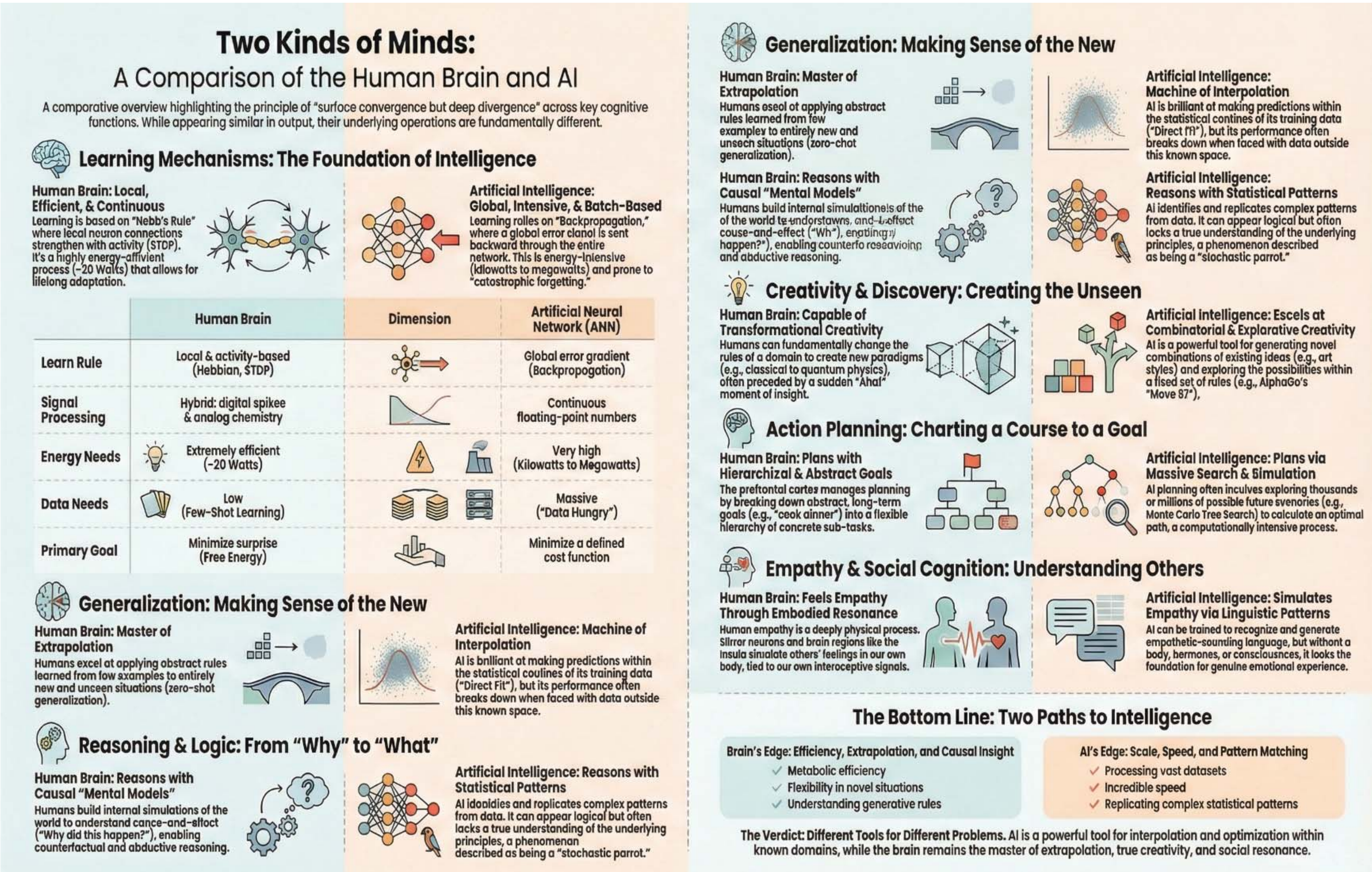
Probably the most critical area is empathy. We increasingly interact with chatbots that sound "understanding." Yet here lurks a danger. Empathy is not a calculation task. In humans, empathy is "embodied." (Rizzolatti & Craighero, 2004; Barrett, 2017; Craig, 2009; Critchley & Garfinkel, 2017; Bender, E. M., et al. 2021)

- **Mirror Neurons:** When we see that someone is in pain, similar areas in our brain fire as if we were in pain ourselves. We simulate the other *in our own body*. (Rizzolatti & Craighero, 2004)
- **Interoception:** We use our gut feeling, our heartbeat, our own physique to feel. (Craig, 2009; Critchley & Garfinkel, 2017; Barrett, 2017)

The AI lacks this body. It has no hormones, no pain, no fear of death. When an AI says: "I understand that you are sad," this is a linguistic simulation. They are vectors in a high-dimensional space that statistically likely follow your input. There is no one there who feels. The *phenomenology* (the subjective experience) is missing. The use of "empathic AI" in therapy or care therefore carries the risk of a gigantic deception: We project feelings onto something that is mathematically incapable of reciprocating them. (Barrett, 2017; Bickmore & Picard, 2005; Fitzpatrick et al., 2017)

Conclusion: Alien Intelligence

The analysis clearly shows: We should stop measuring AI by the human standard. It is not a worse version of a brain and also not a better one. It is an "Alien Intelligence." It optimizes instead of understanding. It interpolates instead of extrapolating. It simulates instead of feeling. That makes it a powerful tool of efficiency, but the domains of true insight, radical creativity, and social resonance remain—for now—our biological privilege. (Hasson et al., 2020; Marcus, 2018; Shanahan, 2022)



References

Attwell, D., & Laughlin, S. B. (2001). An energy budget for signaling in the grey matter of the brain. *Journal of Cerebral Blood Flow & Metabolism*, *21*(10), 1133–1145. https://doi.org/10.1097/00004647-200110000-00001

Barrett, L. F. (2017). *How emotions are made: The secret life of the brain*. Houghton Mifflin Harcourt.

Bi, G.-Q., & Poo, M.-M. (1998). Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. *The Journal of Neuroscience*, *18*(24), 10464–10472. https://doi.org/10.1523/JNEUROSCI.18-24-10464.1998

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. https://doi.org/10.1145/3442188.3445922

Bickmore, T. W., & Picard, R. W. (2005). Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction*, *12*(2), 293–327. https://doi.org/10.1145/1067860.1067867

Boden, M. A. (2004). *The creative mind: Myths and mechanisms (2nd ed.)*. Routledge.

Chollet, F. (2019). *On the measure of intelligence*. arXiv. https://doi.org/10.48550/arXiv.1911.01547

Craig, A. D. (2009). How do you feel—now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, *10*(1), 59–70. https://doi.org/10.1038/nrn2555

Critchley, H. D., & Garfinkel, S. N. (2017). Interoception and emotion. *Current Opinion in Psychology*, *17*, 7–14. https://doi.org/10.1016/j.copsyc.2017.04.020

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, *114*(4), 864–886. https://doi.org/10.1037/0033-295X.114.4.864

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. *JMIR Mental Health*, *4*(2), e19. https://doi.org/10.2196/mental.7785

Geirhos, R., Jacobsen, J.-H., Michaelis, C., Zemel, R., Brendel, W., Bethge, M., & Wichmann, F. A. (2020). Shortcut learning in deep neural networks. *Nature Machine Intelligence*, *2*(11), 665–673. https://doi.org/10.1038/s42256-020-00257-z

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, *315*(5812), 619. https://doi.org/10.1126/science.1134475

Hasson, U., Nastase, S. A., & Goldstein, A. (2020). Direct fit to nature: An evolutionary perspective on biological and artificial neural networks. *Neuron*, *105*(3), 416–434. https://doi.org/10.1016/j.neuron.2019.12.002

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. Wiley.

Hendrycks, D., & Dietterich, T. G. (2019). *Benchmarking neural network robustness to common corruptions and perturbations (arXiv:1903.12261).* arXiv. https://arxiv.org/abs/1903.12261

Jung-Beeman, M., Bowden, E. M., Haberman, J., Frymiare, J. L., Arambel-Liu, S., Greenblatt, R., Reber, P. J., & Kounios, J. (2004). Neural activity when people solve verbal problems with insight. *PLOS Biology*, *2*(4), e97. https://doi.org/10.1371/journal.pbio.0020097

Kounios, J., & Beeman, M. (2009). The Aha! moment: The cognitive neuroscience of insight. *Current Directions in Psychological Science*, *18*(4), 210–216. https://doi.org/10.1111/j.1467-8721.2009.01638.x

Lake, B. M., Salakhutdinov, R., & Tenenbaum, J. B. (2015). Human-level concept learning through probabilistic program induction. *Science*, *350*(6266), 1332–1338. https://doi.org/10.1126/science.aab3050

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences*, *40*, e253. https://doi.org/10.1017/S0140525X16001837

Levy, W. B., & Calvert, V. G. (2021). Communication consumes 35 times more energy than computation in the human cortex, but both costs are needed to predict synapse number. *Proceedings of the National Academy of Sciences*, *118*(18), e2008173118. https://doi.org/10.1073/pnas.2008173118

Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., & Hinton, G. (2020). Backpropagation and the brain. *Nature Reviews Neuroscience*, *21*(6), 335–346. https://doi.org/10.1038/s41583-020-0277-3

Marcus, G. (2018). *Deep learning: A critical appraisal (arXiv:1801.00631).* arXiv. https://arxiv.org/abs/1801.00631

Patterson, D., Gonzalez, J., Le, Q., Liang, C., Munguia, L.-M., Rothchild, D., So, D., Texier, M., & Dean, J. (2021). *Carbon emissions and large neural network training (arXiv:2104.10350).* arXiv. https://arxiv.org/abs/2104.10350

Pearl, J., & Mackenzie, D. (2018). *The book of why: The new science of cause and effect*. Basic Books.

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge University Press.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–192. https://doi.org/10.1146/annurev.neuro.27.070203.144230

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533–536. https://doi.org/10.1038/323533a0

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, *80*(1), 1–27. https://doi.org/10.1152/jn.1998.80.1.1

Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. *Communications of the ACM*, *63*(12), 54–63. https://doi.org/10.1145/3381831

Shanahan, M. (2022). *Talking about large language models (arXiv:2212.03551).* arXiv. https://arxiv.org/abs/2212.03551

Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 3645–3650). Association for Computational Linguistics. https://doi.org/10.18653/v1/P19-1355

Waytz, A., Cacioppo, J., & Epley, N. (2010). Who sees human? The stability and importance of individual differences in anthropomorphism. *Perspectives on Psychological Science*, *5*(3), 219–232. https://doi.org/10.1177/1745691610369336

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. https://doi.org/10.1145/365153.365168