Experimental Evidence for the Propagation and Preservation of Machine Discoveries in Human Populations

Levin Brinkmann[1†], Thomas F. Eisenmann[1*†], Anne-Marie Nussberger[1†], Maxim Derex[2], Sara Bonati[1], Valerii Chirkov[3], Iyad Rahwan[1*]

[1]Center for Humans and Machines, Max Planck Institute for Human Development, Berlin 14195, Germany.

[2]Institute for Advanced Study in Toulouse, Toulouse School of Economics, Toulouse 31080, France.

[3]Institute for Theoretical Biology, Humboldt University of Berlin, Berlin 10099, Germany.

[*]Corresponding authors. Email: mueller@mpib-berlin.mpg.de (T.F.E.); rahwan@mpib-berlin.mpg.de (I.R.)

[†]These authors contributed equally to this work.

Intelligent machines with superhuman capabilities have the potential to uncover problem-solving strategies beyond human discovery. Emerging evidence from competitive gameplay, such as Go, demonstrates that AI systems are evolving from mere tools to sources of cultural innovation adopted by humans. However, the conditions under which intelligent machines transition from tools to drivers of persistent cultural change remain unclear. We identify three key conditions for machines to fundamentally influence human problem-solving: the discovered strategies must be non-trivial, learnable, and offer a clear advantage. Using a cultural transmission experiment and an agent-based simulation, we demonstrate that when these conditions are met, machine-discovered strategies can be transmitted, understood, and preserved by human populations, leading to enduring cultural shifts. These findings provide a framework for understanding how machines can persistently expand human cognitive skills and underscore the need to consider their broader implications for human cognition and cultural evolution.

Machines powered by Artificial Intelligence (AI) have the potential to discover strategies that are difficult for humans to conceive: their computational capacity exceeds human processing power and speed; their ability to process and share information in serial or parallel with high fidelity allows them to distribute problem-solving efficiently (*1*). AlphaGo, an AI system that defeated world champion Lee Sedol in the ancient game of Go, discovered superhuman strategies by playing millions of games against a copy of itself, and integrated these experiences into a complex strategy encoded in its neural network (*2*). This strategy has been described as 'alien' considering thousands of years of human Go-play (*3*). Recent evidence suggests that humans are beginning to change their own Go-play as a result of this innovation (*4*). The domain of gameplay remains, however, a rare example of human problem-solving strategies being shaped by machine-discovered strategies. This raises a question: Beyond being mere tools, under which conditions can machines persistently reshape the way we solve problems?

We address this question by integrating insights from the field of Cultural Evolution (*5*): culture changes through a process of variation, transmission, and selection (*6*). We propose that, correspondingly, machine-induced cultural shifts are contingent on the convergence of three critical conditions. *Appropriate Discovery Difficulty:* Strategies must be non-trivial and extend beyond the range of variations typically discovered by humans. *Low Transmission Difficulty:* Humans must be able to learn and effectively disseminate machine-introduced discoveries. *Recognizable Selective Advantage:* Cultural adoption of machine discoveries hinges on their demonstrated success and the extent to which humans attend to them and endorse them.

In more detail, for a new problem-solving strategy to become part of the human cultural repertoire, it first needs to be discovered. Human discovery is limited by computational capacities and pervasive cognitive biases (*1, 7, 8*). For example, human planning is generally confined to a few steps due to the combinatorial explosion of possible actions and outcomes (*9*). In adapting to this challenge, humans use a range of heuristics (*7*), such as prioritizing immediate over long-term outcomes (*10*) or favoring familiar options (*11*). In contrast, machines surpass humans in computational capacity, allowing them to rapidly accumulate experiences that would be too costly in terms of time, computation, or risk for humans to access. This may lead machines to discover strategies that are counterintuitive to humans yet have superior performance (see Figure 1B, vertical axis).

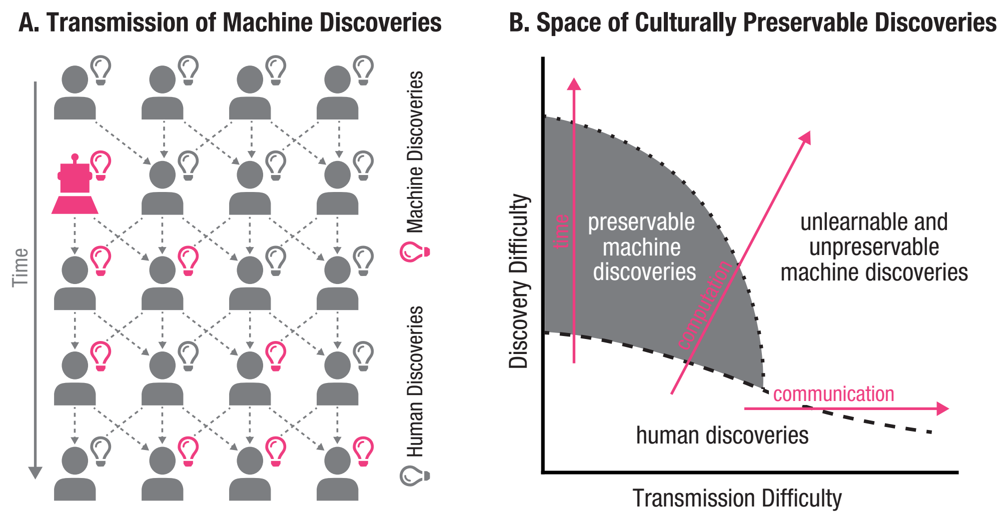

Figure 1: **A. Transmission of machine discoveries can lead to cultural shifts.** The multitude of potential transmission paths enables populations to preserve strategies that are difficult to transmit, fostering resilience against cultural loss. Machines, with their unique capabilities, may discover strategies that are improbable—or difficult to conceive—for humans. If such machine-discovered strategies are successfully transmitted and preserved within human populations, they could drive cultural shifts that outlast the presence of machines. **B. Space of culturally preservable strategies.** Human populations adopt strategies that are either easy to learn or sufficiently transmittable, forming a cultural frontier shaped by human limitations in time, computation, and communication (dashed line). We propose that some strategies, while practically inconceivable, may still be preservable by humans (shaded area). Machines, by discovering these strategies, could shift the cultural frontier (dashed line). Conversely, some machine-discovered strategies might not be preservable by humans, potentially fostering a reliance on machines (outer space).

Discovering superior strategies is not sufficient to induce a cultural shift, however (*12*). The superior performance of the machine-discovered strategy must also be easy for humans to recognize as such (*13–16*). Furthermore, humans must be able to socially learn the strategy and transmit it to others (*6*). The ability to select superior strategies and to transmit them in combination enables the population to maintain complex strategies that would be beyond individuals' reach (*17, 18*). When these mechanisms apply to machine-discovered strategies as well, machines could become sources of new knowledge that shift human culture (see Figure 1B, horizontal axis).

Even though intelligent machines are increasingly used in domains of human discovery, such as education and research (*19–21*), in most cases these AI systems do not convey problem-solving strategies that are entirely alien to human culture. As such, they do not satisfy the above conditions but remain within the confines of existing human discoveries (see Figure 1B bottom left). And even those AI systems that *have* discovered super-human knowledge, remain –at present– tools in human hands rather than integrated into human cognitive frameworks. AlphaFold provides a remarkable example for this category: while it was inspired by the success of AlphaGo in emulating the intuition of human experts (*22*) and in doing so has achieved solving protein-folding problems at superhuman levels (*23, 24*), its discoveries have not (yet) transpired into novel human intuitions, unlike in the case of AlphaGo. As such, it presents an unlearnable and unpreservable machine discovery (see Figure 1B top right).

To provide the first experimental evidence of preserved, machine-induced cultural shift in humans, we designed an experimental setup that features conditions of appropriate discovery and transmission difficulty, along with a recognizable selective advantage for machine performance. The task emulated a prototypical situation in which human exploration is constrained by the pervasive human tendency to avoid losses (i.e., myopic decision-making). Human participants (n = 1,155) were paid to solve a series of "reward networks" (Figure 2 C), which we developed as a modified version of an established paradigm (*25, 26*). In our task, networks consisted of nodes connected by links that were associated with rewards ranging from -50 to +400 points. Participants' goal was to choose 10 moves through the network that would maximize the number of points gained. Our networks were designed such that there was an optimal strategy for solving them: incurring early losses subsequently led to links associated with the maximum reward of +400 points, which were not accessible when avoiding early losses. Hence, the optimal strategy for solving our networks contradicts the myopic tendency of humans to avoid losses and prioritize immediate gains, even when accepting short-term losses could be beneficial in the long run. In line with prior literature, we expected that this myopic tendency would make the discovery of optimal strategy in our task difficult for human players (*27*). Meanwhile, we expected that a few minutes of training on our task would be sufficient for a deep-Q-learning algorithm (the "machine player") to successfully discover the optimal strategy. To compare the performance of human players on their own against the performance of human players exposed to machine players, we randomly allocated participants into

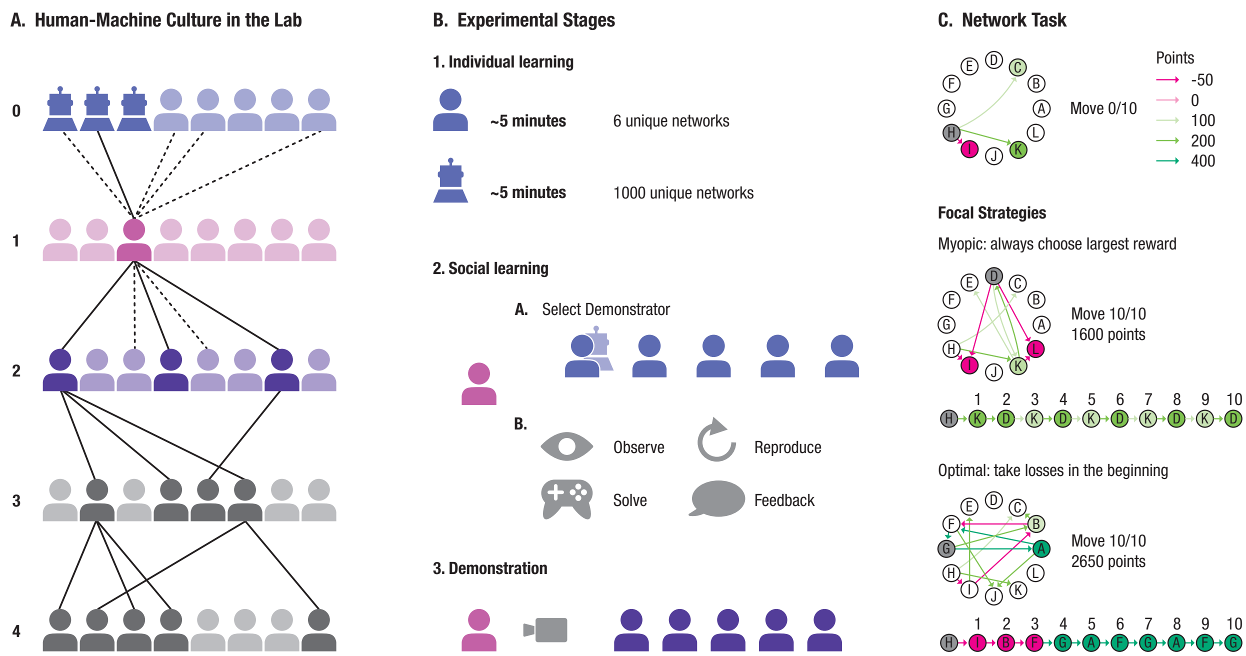

Figure 2: **A. Human-Machine Culture in the Lab.** Participants were divided into populations of five Generations (0-4) with eight individuals each, including 15 Human-Machine Populations (three machines in Generation 0) and 15 Human-Only Populations. Lines represent a subset of demonstrator-learner connections (dashed: potential, solid: selected) involving a focal participant (dark magenta) in Generation 1 and its cultural ancestor (dark blue) and descendants (dark violet and grey). **B. Experimental Stages.** The 'Individual Learning Phase' lasted five minutes during which humans in Generation 0 managed to train on six networks, while machines managed to train on 1,000 networks. In the 'Social Learning Phase', Generations 1-4 were informed about demonstrator candidates' performance and selected their preferred demonstrator. They then observed and replicated the demonstrator's strategy before solving the same network themselves and receiving feedback. In the 'Demonstration Phase,' individuals recorded their strategy for the next generation. **C. Network Task.** Participants collected points on a network, with new potential connections appearing sequentially. We illustrate the two focal strategies in our task: maximizing immediate gains (myopic), and incurring initial costs for greater future rewards (optimal).

populations that involved either only eight humans (15 "Human-Only Populations") or five humans and three machine players in their initial generation (15 "Human-Machine Populations"). Each population consisted of five generations (numbered 0-4; Fig. 2 A), allowing us to observe differences between the two kinds of populations over time. After a series of individual attempts at solving networks, participants in Generation 0 solved four networks during the ”Demonstration Phase”, knowing that their moves would be recorded for future generations to learn from. Participants from Generations 1-4 could choose a ”demonstrator” from the previous generation to learn from, based on the demonstrators' average performance. Only after this Social Learning Phase did participants from Generations 1-4 demonstrate their moves for the next generation. All networks solved by the participants throughout the experiment were unique, requiring them to discover or learn a generalizable strategy rather than optimizing a specific solution.

In a second exploratory step, we sought to follow up on our experimental results with an agent-based simulation. Removing the feasibility constraints that apply to large-scale experiments with human participants, the simulation allowed us to systematically vary discovery and transmission difficulty, and to pitch random and selective social learning against one another (*28*). In parallel to the experimental set-up, we organized agents into five generations, each consisting of eight agents. Agents could discover strategies through exploration or adopt strategies from the previous generation. Abstracting away from the task we used in the experiment, agents explored a trinary strategy space comprising the increasingly rewarding strategies: *random*, *myopic*, and *optimal*. For human-machine populations, three ”machine agents” with a discovery rate 1000 times higher than that of ”human agents” were placed in the first generation.

Humans Preserve the Machine Discovery

In our reward network task, achieving more than 2000 points was only possible by following the optimal strategy of incurring early losses to reach later links associated with the maximum reward of +400 points per move. Generally, human players did not find this optimal strategy on their own: Among the human populations, only one out of 600 participants *discovered* the optimal strategy over the course of the experiment, which four other human players then learned from them. In contrast, we found that the optimal strategy readily spread from machines to humans in the human-machine populations, resulting in the machine-discovery being preserved beyond the presence of the machines and across multiple generations: In 9 out of 15 Human-Machine Populations, human players learned this strategy in each of their generations. Correspondingly, we classified these populations as *permanently preserving* the machine discovery. In the remaining 6 populations, at least one human player in the generation immediately following the machine achieved more than 2000 points; but this discovery was only *temporarily preserved* as it did not transmit to later generations. Figure 3 illustrates prototypical cases for each of these qualitative categories. To summarise, the majority of Human-Machine Populations preserved the optimal strategy until the final generation in our experiment, while 14 out of 15 human populations did *not discover* the optimal strategy at all.

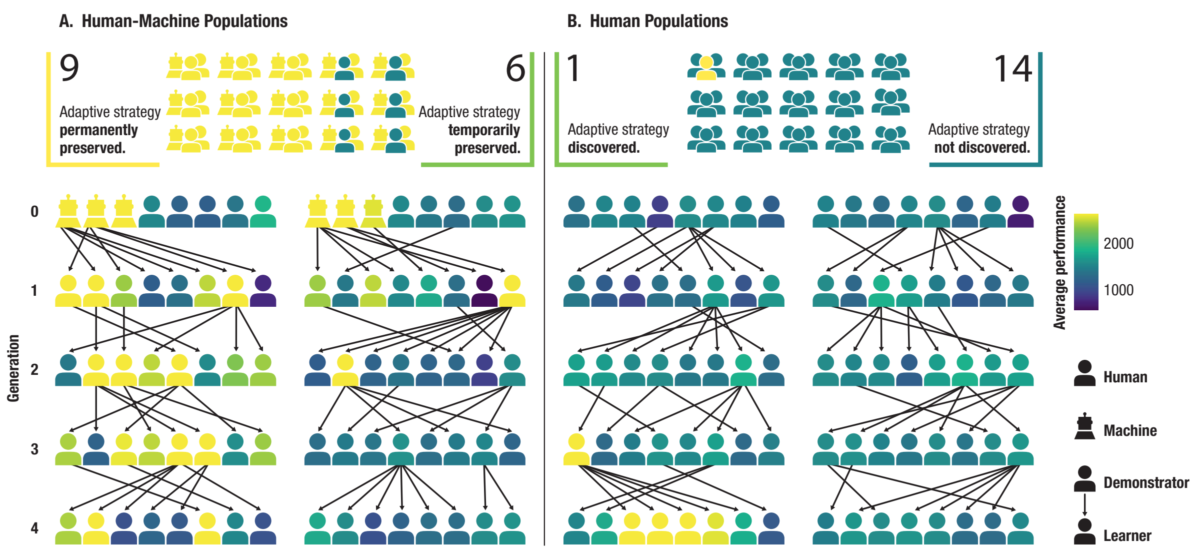

Figure 3: **Four prototypical populations**. We categorized populations based on the average scores of their top performers across generations. Scores above 2000 points indicate adoption of the optimal strategy. Among 15 Human-Machine Populations (panel A), 9 consistently exceeded this threshold, while the remaining 6 surpassed it temporarily in at least one generation following the introduction of machines. In contrast, only 1 of 15 Human-Only Populations (panel B) temporarily achieved scores indicative of the optimal strategy, while all the other Human-Only Populations did not. We illustrate one prototypical population for each case, with colors reflecting performance scores. Arrows represent demonstrator-learner relationships.

In terms of cultural lineages that emerged from demonstrator-learner relationships among the Human-Machine Populations (see Figure 3), we found that only machines had long-lasting impact on the cultural dynamics operating across generations: In the first socially learning generation of the Human-Machine Populations, most players (90%) selected a machine as their demonstrator, and by the last generation, every player was a descendant of a machine player. In contrast, none of the human players in Generation 0 of the Human-Machine Populations initiated a cultural lineage that persisted until the last generation.

Humans Benefit from the Machine Discovery

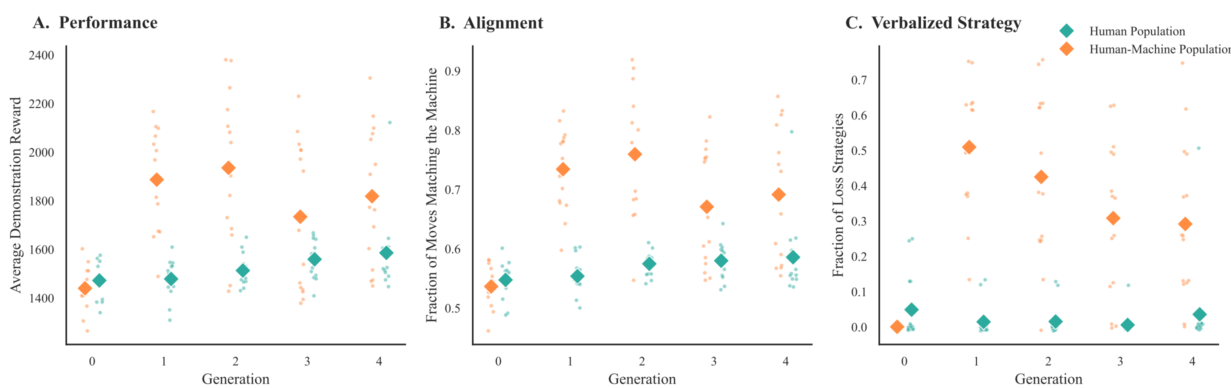

Figure 4: **Evolution of Task Performance, Machine Alignment, and Strategy Description**. Individual populations are represented as dots, with the grand average for each condition indicated by a diamond symbol. Human-Only Populations are indicated in green, Human-Machine Populations in orange. Performance (Panel A) refers to the average number of points earned during the Demonstration Phase. Alignment (Panel B) indicates the proportion of human player moves that a machine player would have made as well (contingent on making the same previous moves). Verbalised Strategy (Panel C) shows the fraction of human players who articulated a strategy referring to a deliberate acceptance of losses after the Social Learning Phase. Data points are jittered horizontally and vertically to improve visibility.

Panel A in Figure 4 shows the average reward gained in the task by generation and population. Descriptively, while the two types of populations started out with the same average in Generation 0, in Generations 1-4 Human-Machine Populations consistently earned higher rewards. Our pre-registered linear mixed-effects model on these latter generations supported that this difference was significant (β = −309.2, SE = 57.0, $CI_{95\%}$ = [−420.7, −196.6]). Further models confirmed that Human-Machine Populations performed better than Human-Only Populations both in Generation

1, which could learn directly from machine players ($\beta = -404.6$, $SE = 59.4$, $\mathrm{CI}_{95\%} = [-513.5, -293.0]$), and in Generation 4, where influence from machines was remote by four iterations ($\beta = -229.1$, $SE = 82.4$, $\mathrm{CI}_{95\%} = [-407.4, -72.3]$).

Humans Align with the Machines

Panel B in Figure 4 shows the proportion of human moves aligned with machine-generated moves, summarized by generation and population. Qualitatively, the results are similar to the results for rewards: no difference between conditions in Generation 0, but Human-Machine Populations are consistently more aligned in Generations 1 to 4. A logistic mixed-effects model on these generations shows that Human-Machine Populations were more machine-aligned than human populations ($\beta = -0.88$, $SE = 0.14$, $\mathrm{CI}_{95\%} = [-1.14, -0.54]$). Again, this was also the case within Generation 1, which learned from machine players ($\beta = -1.16$, $SE = 0.14$, $\mathrm{CI}_{95\%} = [-1.43, -0.90]$), and Generation 4 ($\beta = -0.66$, $SE = 0.19$, $\mathrm{CI}_{95\%} = [-1.03, -0.24]$).

Humans Internalize the Machine Discovery

During the Demonstration Phase of our experiment, players were asked to provide a written strategy of how they approached the task. In contrast to the demonstration itself, these written strategies were not transmitted to the next generation, and players were aware of this. Panel C in Figure 4 shows the proportion of human players mentioning the optimal strategy (i.e. the ”loss strategy”) in their written statements, as coded independently by three human annotators. In the Human-Only Populations, we observe practically no mention of the strategy, with proportions close to 0 throughout all generations; the exception to this is the single population that discovered the optimal strategy in Generation 3, resulting in 50% of players referring to it in Generation 4. In contrast, participants in the Human-Machine Populations consistently mentioned the optimal strategy explicitly, with proportions of about 30% to 50% per generation. A logistic mixed-effects model on Generations 1 to 4 shows that this difference is significant ($\beta = -4.18$, $SE = 0.62$, $\mathrm{CI}_{95\%} = [-5.49, -3.04]$).

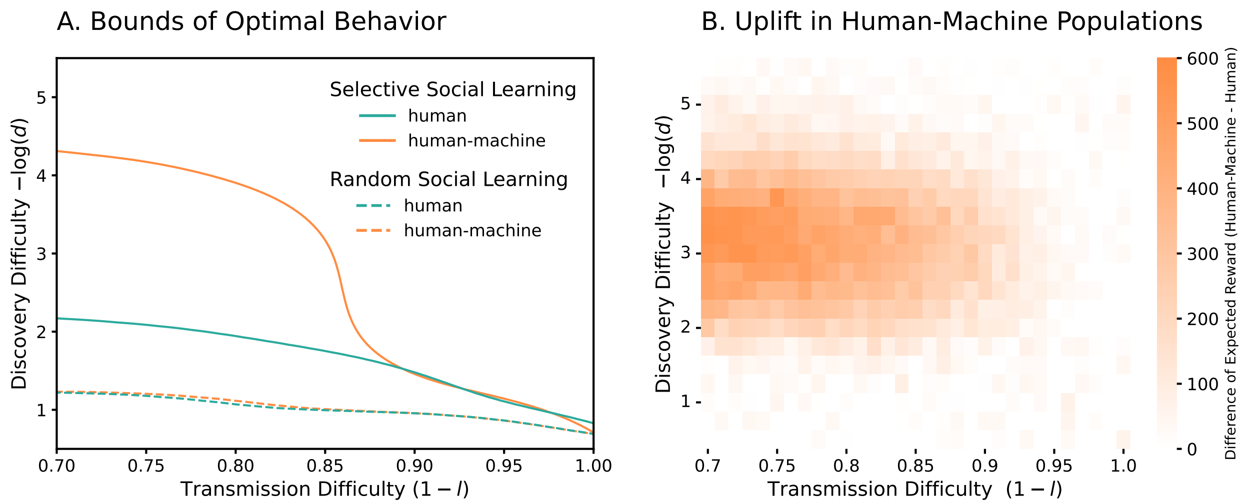

Figure 5: **Agent-Based Simulation Highlights Dependence on Discovery and Transmission Rates.** An agent-based simulation replicates our five-generation social learning experiment, varying the difficulty of discovering and transmitting the optimal strategy. Panel A shows the 50% adoption boundary for the optimal strategy in the final generation, with lower discovery and transmission difficulty promoting its adoption, particularly under selective social learning (solid lines). Machines enable the adoption of hard-to-conceive but transmittable optimal strategies (orange line). Panel B illustrates the boost in average reward of mixed populations over human-only populations in the final generation, highlighting that this persistent boost is limited to strategies that are challenging for humans to discover but accessible to machines and that are of moderate transmission difficulty. Here we show the uplift for agents with selective social learning only.

Simulation Corroborates Conditions for Cultural Shift

We devised an agent-based simulation to systematically explore how variations in discovery and transmission difficulty shape the adoption of machine-discoveries and, ultimately, population-level success. Corroborating our experimental results, harder-to-discover strategies (with a higher (log(d)) were adopted when transmission difficulty $1 - t$ was low, where t is the transmission rate towards the end of the multi-generational simulation (Panel A in Fig. 5). This effect was enhanced when agents selected the highest-performing individual from the previous generation to socially learn from (solid lines) compared to learning from a random agent (dashed lines). When machine agents, with their increased discovery rate, were added, these human-machine populations could discover and preserve the optimal solution even when discovery was difficult (orange solid line). This machine-induced effect was not present when agents learned from a random agent of the previous generation (orange dashed line).

The ability of machine agents to discover hard-to-conceive strategies, combined with the ability of human agents to adopt and preserve these strategies, led to an increase in the rewards collected by human agents in the last generation of human-machine populations compared to their counterparts in human-only populations (Panel B in Fig. 5). We found that this uplift was constrained by both (a) high discovery difficulty, which renders human discoveries unlikely but machine discoveries likely, and (b) moderate transmission difficulty, which makes transmission feasible.

Discussion

Our work demonstrates that machine-discoveries can persistently reshape human problem-solving strategies when they are calibrated in difficulty, transmissibility, and selective advantage. In our experiment, superhuman computational capacity allowed machine players to discover a superior strategy that offset human cognitive biases toward myopic and suboptimal solutions and, as such, was effectively beyond human reach. Yet, the machine-discovered strategies were still learnable for human players, who did not merely replicate them but also applied them to new variations of the task. Further signifying the integration of the alien problem-solving strategy into their cognitive framework, many human players explicitly articulated the strategy in writing. The machine-discovery's selective advantage translated into human players overwhelmingly favoring machine over human ancestors, producing persistent cultural lineages where every human player in the last generation was ultimately a descendant of a machine player in the first generation.

A major question arising from these results concerns the scope of machine-induced cultural shifts. Our pilot experiments illustrated the key challenge of achieving the right balance between discovery difficulty and transmissibility: At one extreme, simple strategies were readily discovered by humans without machine assistance, limiting their capacity to induce meaningful shifts. At the other extreme, strategies requiring lengthy and precise sequences of actions were too complex for human participants to learn and pass on. Yet, prior work from cultural evolution has shown that large effective population size can effectively buffer against risks of losing complex discoveries (*29, 30*).

This suggests that the scope of machine-induced cultural shifts may expand in contexts where large groups interact with machine-generated strategies. For instance, the widespread adoption of generative AI in educational, professional, and creative contexts naturally exposes millions of people to machine-mediated strategies for creation or problem-solving. Larger population sizes increase the probability that machine-discovered strategies are learned in the first place, as a diverse range of individuals with varying levels of skills engage with them. Analogously, skilled individuals can prevent strategies from being lost, allowing less skilled individuals to benefit from more demonstrations and eventually acquire the strategies (*30*), thereby ultimately accumulating a critical mass of individuals for triggering a tipping point in their cultural adoption (*31*). Emergent evidence suggests that even simple machine agents could further boost these group dynamics' efficacy (*32*). Yet another factor that could broaden the scope of machine-induced cultural shifts is pedagogical behavior. Teaching is one of the primary mechanisms through which humans transmit knowledge and skills (*6, 33*), and experiments have shown that the value of teaching increases with the complexity of strategies (*34*). Notably, discovery and transmission difficulty have also been identified as key determinants in the evolution of teaching (*35*), underscoring that machine-induced cultural shifts may occur within a relatively narrow range of conditions. This line of work does, however, also point towards another possible pathway that could enhance the efficacy of machine-induced cultural shift: more explainable AI systems could provide both teachers and learners with insights that facilitate the adoption of machine-discovered strategies, and enhance their scope through more effective human-machine interactions (*36*).

The work presented here suggests that by learning from machines, humans may adopt and transmit hard-to-conceive optimal strategies, thereby broadening humans' cognitive repertoire without incurring the costs typically associated with discovering such strategies. While our findings are primarily limited to scenarios in which human behavior is suboptimal due to cognitive constraints, future research could explore how machine-discovered strategies might also mitigate exploration limitations imposed by structural factors, such as entrenched path dependencies or environmental pressures towards predictability in contexts such as scientific discovery (*21, 37*). Supporting this conjecture, recent work has highlighted machines' potential to identify novel strategies for fostering cooperation that humans have proven unlikely to consider (*38*). Even when it comes to changing notoriously entrenched conspiracy beliefs, intelligent machines (here, LLMs) are emerging as potent agents of persistent learning among humans (*39*). But these contexts also point to the potential challenge that machine-discovered strategies might sometimes be undesirable from a human perspective and not be explored for good reasons, such as ethical concerns or social acceptability. Relatedly, human adoption and preservation of machine-discovered strategies may raise ethical concerns about human autonomy and creativity as much as it might bear epistemic risks that could undermine human knowledge (*40*). And yet, these very concerns underscore the transformative potential of machines pushing the boundaries of human cognition, enabling the discovery of hard-to-conceive strategies and ultimately facilitating cultural accumulation.

References and Notes

1. T. L. Griffiths, Understanding Human Intelligence through Human Limitations. *Trends in Cognitive Sciences* **24** (11), 873–883 (2020), doi:10.1016/j.tics.2020.09.001, https://www.sciencedirect.com/science/article/pii/S1364661320302151.

2. D. Silver, *et al.*, Mastering the game of Go with deep neural networks and tree search. *Nature* **529** (7587), 484–489 (2016), publisher: Nature Publishing Group, doi:10.1038/nature16961, https://www.nature.com/articles/nature16961.

3. C. Sang-Hun, Google's Computer Program Beats Lee Se-dol in Go Tournament. *The New York Times* (2016), http://www-nytimes-com/2016/03/16/world/asia/korea-alphago-vs-lee-sedol-go.html.

4. M. Shin, J. Kim, B. van Opheusden, T. L. Griffiths, Superhuman artificial intelligence can improve human decision-making by increasing novelty. *Proceedings of the National Academy of Sciences* **120** (12), e2214840120 (2023), publisher: Proceedings of the National Academy of Sciences, doi:10.1073/pnas.2214840120, http://www-pnas-org/doi/10.1073/pnas.2214840120.

5. L. Brinkmann, *et al.*, Machine culture. *Nature Human Behaviour* **7** (11), 1855–1868 (2023), publisher: Nature Publishing Group, doi:10.1038/s41562-023-01742-2, https://www.nature.com/articles/s41562-023-01742-2.

6. R. Boyd, P. J. Richerson, *Culture and the evolutionary process*, Culture and the evolutionary process (University of Chicago Press, Chicago, IL, US) (1985), pages: viii, 331.

7. H. A. Simon, A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics* **69** (1), 99–118 (1955), publisher: Oxford University Press, doi:10.2307/1884852, https://www.jstor.org/stable/1884852.

8. A. Tversky, D. Kahneman, Judgment under Uncertainty: Heuristics and Biases. *Science* **185** (4157), 1124–1131 (1974), publisher: American Association for the Advancement of Science, doi:10.1126/science.185.4157.1124, http://www-science-org/doi/abs/10.1126/science.185.4157.1124.

9. H. A. Simon, A. Newell, Human problem solving: The state of the theory in 1970. *American Psychologist* **26** (2), 145–159 (1971), place: US Publisher: American Psychological Association, doi:10.1037/h0030806.

10. S. Frederick, G. Loewenstein, T. O'Donoghue, Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature* **40** (2), 351–401 (2002), doi:10.1257/002205102320161311, `http://www-aeaweb-org/articles?id=10.1257/002205102320161311`.

11. D. G. Goldstein, G. Gigerenzer, Models of ecological rationality: The recognition heuristic. *Psychological Review* **109** (1), 75–90 (2002), place: US Publisher: American Psychological Association, doi:10.1037/0033-295X.109.1.75.

12. I. Rahwan, D. Krasnoshtan, A. Shariff, J.-F. Bonnefon, Analytical reasoning task reveals limits of social learning in networks. *Journal of The Royal Society Interface* **11** (93), 20131211 (2014), publisher: Royal Society, doi:10.1098/rsif.2013.1211, `https://royalsocietypublishing.org/doi/abs/10.1098/rsif.2013.1211`.

13. J. Henrich, F. J. Gil-White, The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior* **22** (3), 165–196 (2001), doi:10.1016/S1090-5138(00)00071-4, `https://www.sciencedirect.com/science/article/pii/S1090513800000714`.

14. R. L. Kendal, I. Coolen, Y. van Bergen, K. N. Laland, Trade-Offs in the Adaptive Use of Social and Asocial Learning, in *Advances in the Study of Behavior* (Academic Press), vol. 35, pp. 333–379 (2005), doi:10.1016/S0065-3454(05)35008-X, `https://www.sciencedirect.com/science/article/pii/S006534540535008X`.

15. R. McElreath, *et al.*, Beyond existence and aiming outside the laboratory: estimating frequency-dependent and pay-off-biased social learning strategies. *Philosophical Transactions of the Royal Society B: Biological Sciences* **363** (1509), 3515–3528 (2008), publisher: Royal Society, doi:10.1098/rstb.2008.0131, `https://royalsocietypublishing.org/doi/abs/10.1098/rstb.2008.0131`.

16. T. J. H. Morgan, L. E. Rendell, M. Ehn, W. Hoppitt, K. N. Laland, The evolutionary basis of human social learning. *Proceedings of the Royal Society B: Biological Sciences* **279** (1729), 653–662 (2011), publisher: Royal Society, doi:10.1098/rspb.2011.1172, `https://royalsocietypublishing.org/doi/abs/10.1098/rspb.2011.1172`.

17. B. Thompson, B. van Opheusden, T. Sumers, T. L. Griffiths, Complex cognitive algorithms preserved by selective social learning in experimental populations. *Science* **376** (6588), 95–98 (2022), publisher: American Association for the Advancement of Science, doi:10.1126/science.abn0915, `http://www-science-org/doi/10.1126/science.abn0915`.

18. E. Miu, N. Gulley, K. N. Laland, L. Rendell, Innovation and cumulative culture through tweaks and leaps in online programming contests. *Nature Communications* **9** (1), 2321 (2018), publisher: Nature Publishing Group, doi:10.1038/s41467-018-04494-0, `https://www.nature.com/articles/s41467-018-04494-0`.

19. X. Luo, *et al.*, Large language models surpass human experts in predicting neuroscience results. *Nature Human Behaviour* pp. 1–11 (2024), publisher: Nature Publishing Group, doi:10.1038/s41562-024-02046-9, `https://www.nature.com/articles/s41562-024-02046-9`.

20. B. Sanchez-Lengeling, A. Aspuru-Guzik, Inverse molecular design using machine learning: Generative models for matter engineering. *Science* **361** (6400), 360–365 (2018), publisher: American Association for the Advancement of Science, doi:10.1126/science.aat2663, `http://www-science-org/doi/full/10.1126/science.aat2663`.

21. S. Fortunato, *et al.*, Science of science. *Science (New York, N.Y.)* **359** (6379), eaao0185 (2018), doi:10.1126/science.aao0185, `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5949209/`.

22. W. D. Heaven, This is the reason Demis Hassabis started DeepMind (2022), `https://www.technologyreview.com/2022/02/23/1045016/ai-deepmind-demis-hassabis-alphafold/`.

23. J. Jumper, *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596** (7873), 583–589 (2021), publisher: Nature Publishing Group, doi:10.1038/s41586-021-03819-2, `http://www-nature-com/articles/s41586-021-03819-2`.

24. K. Tunyasuvunakool, *et al.*, Highly accurate protein structure prediction for the human proteome. *Nature* **596** (7873), 590–596 (2021), publisher: Nature Publishing Group, doi:10.1038/s41586-021-03828-1, `http://www-nature-com/articles/s41586-021-03828-1`.

25. Q. J. M. Huys, *et al.*, Bonsai Trees in Your Head: How the Pavlovian System Sculpts Goal-Directed Choices by Pruning Decision Trees. *PLOS Computational Biology* **8** (3), e1002410 (2012), publisher: Public Library of Science, doi:10.1371/journal.pcbi.1002410, `https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1002410`.

26. L. Brinkmann, *et al.*, Hybrid social learning in human-algorithm cultural transmission. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **380** (2227), 20200426 (2022), publisher: Royal Society, doi:10.1098/rsta.2020.0426, `https://royalsocietypublishing.org/doi/10.1098/rsta.2020.0426`.

27. B. Thompson, T. L. Griffiths, Human biases limit cumulative innovation. *Proceedings of the Royal Society B: Biological Sciences* **288** (1946), 20202752 (2021), publisher: Royal Society, doi:10.1098/rspb.2020.2752, `https://royalsocietypublishing.org/doi/10.1098/rspb.2020.2752`.

28. R. L. Kendal, *et al.*, Social Learning Strategies: Bridge-Building between Fields. *Trends in Cognitive Sciences* **22** (7), 651–665 (2018), publisher: Elsevier, doi:10.1016/j.tics.2018.04.003, `http://www-cell-com/trends/cognitive-sciences/abstract/S1364-6613(18)30094-9`.

29. J. Henrich, Demography and Cultural Evolution: How Adaptive Cultural Processes Can Produce Maladaptive Losses—The Tasmanian Case. *American Antiquity* **69** (2), 197–214 (2004), doi:10.2307/4128416, `https://doi.org/10.2307/4128416`.

30. M. Derex, M.-P. Beugin, B. Godelle, M. Raymond, Experimental Evidence for the Influence of Group Size on Cultural Complexity. *Nature* **503** (7476), 389–391 (2013), doi:10.1038/nature12774.

31. D. Centola, J. Becker, D. Brackbill, A. Baronchelli, Experimental evidence for tipping points in social convention. *Science* **360** (6393), 1116–1119 (2018), doi:10.1126/science.aas8827, https://www.science.org/doi/10.1126/science.aas8827.

32. A. Ueshima, M. I. Jones, N. A. Christakis, Simple autonomous agents can enhance creative semantic discovery by human groups. *Nature Communications* **15** (1), 5212 (2024), publisher: Nature Publishing Group, doi:10.1038/s41467-024-49528-y, https://www.nature.com/articles/s41467-024-49528-y.

33. G. Csibra, G. Gergely, Social learning and social cognition: The case for pedagogy, in *Processes of Change in Brain and Cognitive Development: Attention and Performance XXI*, Y. Munakata, M. H. Johnson, Eds. (Oxford University Press), p. 0 (2006), doi:10.1093/oso/9780198568742.003.0011, https://doi.org/10.1093/oso/9780198568742.003.0011.

34. A. J. Lucas, *et al.*, The value of teaching increases with tool complexity in cumulative cultural evolution. *Proceedings of the Royal Society B: Biological Sciences* **287** (1939), 20201885 (2020), publisher: Royal Society, doi:10.1098/rspb.2020.1885, https://royalsocietypublishing.org/doi/10.1098/rspb.2020.1885.

35. L. Fogarty, P. Strimling, K. N. Laland, The Evolution of Teaching. *Evolution* **65** (10), 2760–2770 (2011), doi:10.1111/j.1558-5646.2011.01370.x, https://doi.org/10.1111/j.1558-5646.2011.01370.x.

36. T. W. Malone, R. Laubacher, C. Dellarocas, The Collective Intelligence Genome. *MIT Sloan Management Review* (2010), https://sloanreview.mit.edu/article/the-collective-intelligence-genome/.

37. A. Clauset, D. B. Larremore, R. Sinatra, Data-driven predictions in the science of science. *Science* **355** (6324), 477–480 (2017), publisher: American Association for the Advance-

ment of Science, doi:10.1126/science.aal4217, http://www-science-org/doi/full/10.1126/science.aal4217.

38. K. R. McKee, *et al.*, Scaffolding cooperation in human groups with deep reinforcement learning. *Nature Human Behaviour* **7** (10), 1787–1796 (2023), publisher: Nature Publishing Group, doi:10.1038/s41562-023-01686-7, https://www.nature.com/articles/s41562-023-01686-7.

39. T. H. Costello, G. Pennycook, D. G. Rand, Durably Reducing Conspiracy Beliefs through Dialogues with AI. *Science* **385** (6714), eadq1814 (2024), doi:10.1126/science.adq1814.

40. L. Messeri, M. J. Crockett, Artificial intelligence and illusions of understanding in scientific research. *Nature* **627** (8002), 49–58 (2024), publisher: Nature Publishing Group, doi:10.1038/s41586-024-07146-0, http://www-nature-com/articles/s41586-024-07146-0.

41. Code and data repository for this work (2024), https://doi.org/10.17605/OSF.IO/AF4CB.

42. Experimental interface for this study, https://center-for-humans-and-machines.github.io/reward-network-iii.

43. Preregistration for this study, https://aspredicted.org/3C3_HJH.

44. S. Palan, C. Schitter, Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance* **17**, 22–27 (2018), doi:10.1016/j.jbef.2017.12.004, https://www.sciencedirect.com/science/article/pii/S2214635017300989.

45. V. Mnih, *et al.*, Human-level control through deep reinforcement learning. *Nature* **518** (7540), 529–533 (2015), publisher: Nature Publishing Group, doi:10.1038/nature14236, https://www.nature.com/articles/nature14236.

46. J. Chung, C. Gulcehre, K. Cho, Y. Bengio, Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling (2014), doi:10.48550/arXiv.1412.3555, http://arxiv.org/abs/1412.3555, arXiv:1412.3555.

47. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* **67**, 1–48 (2015), doi:10.18637/jss.v067.i01, https://doi.org/10.18637/jss.v067.i01.

Acknowledgments

We thank Jaeeun Shin and Rodrigo Schettino for their support in data collection. We also thank Peter Dayan and Wolfgang Hönig for their valuable feedback.

Author contributions: L.B., T.F.E., M.D., and I.R. conceived the research project and experimental methodology; L.B., T.F.E., A.-M.N., S.B., and V.C. refined the methodology; L.B., T.F.E., A.-M.N., M.D., and I.R. developed the theoretical grounding; T.F.E. pre-registered and conducted the formal analysis; L.B. conducted the agent-based simulation; S.B. developed the machine player; V.C. implemented the experiment; L.B., T.F.E., and A.-M.N. administered and supervised the study; L.B.,T.F.E., and A.-M.N. created the visualizations; L.B., T.F.E., A.-M.N., M.D., and I.R. wrote the paper.

Competing interests: There are no competing interests to declare.

Data and materials availability: All data and code for reproducing the statistical analyses, visualizations, training of the machine player, and execution of the experiment are available on OSF (*41*), including a static version of the experimental interface (*42*). A preregistration document is available on AsPredicted (*43*).

Supplementary materials

Materials and Methods

Supplementary Text

Figs. S1 to S10

Tables S1 to S3