TOPICS
TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science (2019) 1–12

ISSN:1756-8765 online
DOI: 10.1111/tops.12463

This article is part of the topic “Commentaries on Rafael Núñez’s article, "What happened to cognitive science?,” Wayne D. Gray (Topic Editor). For a full listing of topic papers, see http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1756-8765/earlyview

Becoming Cognitive Science

Robert L. Goldstone

Department of Psychological and Brain Sciences, Indiana University

Received 5 August 2019; received in revised form 11 September 2019; accepted 11 September 2019

Abstract

Cognitive science continues to make a compelling case for having a coherent, unique, and fundamental subject of inquiry: What is the nature of minds, where do they come from, and how do they work? Central to this inquiry is the notion of agents that have goals, one of which is their own persistence, who use dynamically constructed knowledge to act in the world to achieve those goals. An agentive perspective explains why a special class of systems have a cluster of co-occurring capacities that enable them to exhibit adaptive behavior in a complex environment: perception, attention, memory, representation, planning, and communication. As an intellectual endeavor, cognitive science may not have achieved a hard core of uncontested assumptions that Lakatos (1978) identifies as emblematic of a successful research program, but there are alternative conceptions according to which cognitive science has been successful. First, challenges of the early, core tenet of “Mind as Computation” have helped put cognitive science on a stronger foundation—one that incorporates relations between minds and their environments. Second, even if a full cross-disciplinary theoretic consensus is elusive, cognitive science can inspire distant, deep, and transformative connections between pairs of fields. To be intellectually vital, cognitive science need not resemble a traditional discipline with its associated insularity and unchallenged assumptions. Instead, there is strength and resilience in the diverse perspectives and methods that cognitive science assembles together. This interdisciplinary enterprise is fragile and perhaps inherently unstable, as the looming absorption of cognitive science into psychology shows. Still, for many researchers, the excitement and benefits of triangulating on the nature of minds by integrating diverse cases cannot be secured by a stable discipline with an uncontested core of assumptions.

Correspondence should be sent to Robert L. Goldstone, Department of Psychological and Brain Sciences, Indiana University, 110 1 E. 10 th St., Bloomington, IN 47405, USA. E-mail: rgoldsto@indiana.edu

Keywords: Cognitive science; Interdisciplinarity; Computation; Networks; Agents; Minds; Scientometrics; Sociology

1. Introduction

Whether cognitive science will wither away or weather the turbulent paradigmatic waves of science is a critical issue for readers of this journal, as well as cognitive science in all of its other manifestations: university departments, degree programs, the Cognitive Science Society, journals bearing the name, and researchers who self-identify as cognitive scientists. Perhaps the key attribute of all persisting organizational units, including companies, biological organisms, countries, religious communities, and intellectual fields of inquiry, is that they strive for their continued persistence. Those units that do not so strive cease to exist. While scientometricians can objectively and dispassionately articulate the case that cognitive science has not matured into a coherent disciplinary field since its early origins in the 1970s (van den Besselaar, 2018; Vugteveen, Lenders, & Van den Besselaar, 2014), we who consider ourselves cognitive scientists cannot help but receive this news with emotional reactions ranging across denial, defensiveness, frustration, and dejection.

I concur with many of the empirical bases of support that Núñez et al. (2019) marshal to support their contention that cognitive science has not yet materialized the promise for a "new science of the mind" as envisioned by Gardner (1987). Cognitive science, whether denoted by the journals like *Cognitive Science* or *Topics in Cognitive Science*, by membership in the Cognitive Science Society, or by degree-granting curricula at universities, is heavily lopsided toward psychology (Gentner, 2010). Cognitive science has become more diverse in its coverage of topics over the years. Unfortunately, this has been achieved not by increasing its coverage of anthropology, computer science, or philosophy, but by branching out to cover areas within psychology beyond adult cognition: development, neuroscience, social, and emotion (Goldstone & Leydesdorff, 2006; Leydesdorff & Goldstone, 2014). Given the dense and strong connections of areas within psychology, it is perhaps unsurprising, in retrospect, that cognitive science would evolve by bringing in originally neglected areas within psychology, rather than creating permanent bridges between disciplines committed to very different methods and goals.

While I strongly endorse Núñez et al.'s claim that cognitive science has inadequate participation outside of psychology, their statements such as "authors with psychology affiliations have continued to be consistently overrepresented among the core disciplines —by nearly 4 times their expected share" (p. 5) assume that there is an objective way of determining what each traditional discipline's contribution should be. I do not make this assumption, and certainly do not assume that each of the six disciplines should get equal shares. Moreover, the six fields of artificial intelligence (and computer science more broadly), anthropology, linguistics, neuroscience, philosophy, and psychology do not exhaustively cover the intellectual territory of cognitive science. Important contributions and insights about cognitive science also come from economics, literary studies, biology, physics, gender studies, business, law, and medicine, to name just a few pertinent disciplines. Cognitive science should not limit itself to the six fields articulated in the Sloan Foundation's 1978 report on cognitive science. Intellectually vibrant endeavors can and should be capable of reconfiguring themselves.

2. Diversity of perspective

Cognitive science is too lop-sided toward psychology not because of a deviation from any particular distribution of contributions from traditional disciplines, but rather because its underlying topic of inquiry requires and benefits from different perspectives. I am treating the core inquiry of cognitive science to be: What is the nature of minds, where do they come from, and how do they work? This is the fundamental inquiry that has piqued the interest of most people who call themselves cognitive scientists, and it is not reducible to the fundamental questions of other sciences like physics ("What is the nature of matter and energy?"), biology ("What is the nature of life?"), or psychology ("Why do humans and other animals behave as they do?"). Minds follow different principles than material objects (Pinker, 1997). When Jane gives John a book, Jane no longer has it. By contrast, when Jane gives an idea from the book to John, both have it. We may not yet know what things do or could usefully be said to have minds, but minds appear to be needed when a system needs to exhibit adaptive behavior in a complex environment.

Multiple disciplinary perspectives are needed because otherwise, our understanding of minds is likely to be too tied to their most immediately available human manifestation. Anthropology encourages us to appreciate that typical participants in psychology experiments are WEIRD (Henrich, Heine, & Norenzayan, 2010)—Western, Educated, Industrialized, Rich, and Democratic, and that there are robust cross-cultural differences in perception, reasoning, learning, and memory (Beller & Bender, 2015). Even more radically, artificial intelligence and philosophy offer the possibility that human beings may not always be the pinnacle of adaptive reasoning systems. Triangulating on the nature of minds by considering the commonalities shared by their biological, mechanical, and community forms provides constraints that a single perspective cannot provide on its own. By striving to create a general enough conception of minds to capture their diverse manifestations, cognitive science belongs to the relatively uncommon class of "integrative sciences." This class also includes artificial life—creating software, robotic, and synthetic biological life forms to understand the essential nature of life abstracted beyond its DNA-based forms (Agmon, Gates, Churavy, & Beer, 2016; Bedau, 2003; Horowitz & England, 2017). A third exemplar is complex systems, the general study of systems made up of parts that interact to create emergent, global patterns such as positive and negative feedback loops, oscillations, and recursive branching structures (Mitchell, 2009; Page & Miller, 2007; Thurner, Klimek, & Hanel, 2018). Interestingly, adaptation is a theme for all three of these integrative sciences, and some approaches subsume all three (Maturana & Varela, 1973/1980).

3. The core

Given the broader intellectual ecology of traditional disciplines like philosophy, economics, and psychology, and other cross-cutting and more labile integrative sciences like artificial life, information science, and network science, does cognitive science continue to make a compelling case for having a unique and valuable perspective? My own answer remains enthusiastic and positive, in large part because of the promise of general theories for our enterprise's core questions. These questions include: How can an agent make decisions and perform actions that increase its likelihood of flourishing? How do agents aptly perceive and use patterns in the environment? How do agents create and modify organizations through social learning, communication, cooperation, and competition? The relatively abstract term "agent" is used to remain agnostic about the form that the agents take—neurons in a human brain, bees in a hive, workers in a large company, or mixtures of recommendation modules in an artificial intelligence. The term "agent" is still more restrictive than the more general term "system" and captures the commonality that brains, hives, companies, and artificial intelligences have goals, one of which is their own persistence, and they use their dynamically constructed knowledge to act in the world to achieve those goals. An agentive perspective featured prominently in the origins of cognitive science (Miller, Galanter, & Pribram, 1960; Simon, 1969), and it remains an inferentially powerful concept. If a system is aptly construable as an agent, then one can make many inferences about the kinds of capacities that it will have—devices for gathering information from the environment (senses), strategically determining where to seek additional information (attention), storing and retrieving previously acquired information (memory), transforming the information in new ways that allow relevant patterns to be emphasized (representation), and establishing action sequences informed by the transformed information (planning and action).

An agentive perspective allows insightful generalizations across superficially dissimilar systems. For example, agents must make decisions, and the diffusion model (Ratcliff, 1978) is a rigorously formulated and empirically well-tested mechanism for decision making that has been successfully applied to many kinds of agents. In this model, information continuously accumulates and a relative rule is used to determine whether a threshold amount of evidence has been achieved in order to respond. The diffusion model posits an evidence accumulator that begins at a starting point and continuously accumulates noisy evidence at a specific rate. There are upper and lower boundaries that serve as thresholds for triggering a decision. The model makes predictions for response accuracy and the entire distribution of response times. It has been useful for modeling neurons' firing rates in monkeys (Gold & Shadlen, 2007). In humans, it organizes and predicts results from cognitive tasks such as recognition, memory, categorization, and choice (Forstmann, Ratcliff, & Wagenmakers, 2016). It has also been recruited to explain foraging behavior in acellular slime molds (Latty & Beekman, 2011) and bee hives (Marshall et al., 2011). The same model applies across these levels because all of these agents need to make accurate and fast decisions (Bogacz et al., 2006).

A second example of the interdisciplinary utility of agent-based theories is adversarial problem solving. When it is beneficial for an agent to come up with a resilient and flexible solution to a problem that it faces, a general approach that has had noteworthy successes in computer science, economics, psychology, evolutionary biology, and business has been to recruit or create a competitor who is tasked with taking advantage of limitations that it finds in the agent's solutions. Early examples of this from game theory established equilibrium choice strategies that are rational in the sense that a player deviating from a strategy will not have an expected gain when playing other rational players (von Neumann & Morgenstern, 1944). These game theoretic accounts have been applied with quantitative precision to explain real-world behaviors such as the probabilistic but heavily biased decisions of penalty kickers and goalies to kick/predict shots to different locations in a soccer goal (Chiappori, Levitt, & Groseclose, 2002). In computer science, an effective way to have a game-playing Artificial Intelligence (AI) learn evaluative weights for different board features is to have an AI with one set of weights play against another AI with different weights, preserving and modifying the winning AI's weights (Samuel, 1959). In modern instantiations of this idea such as the deep neural network AlphaZero (Silver et al., 2018), adversarial self-play among variants allows the same algorithm to produce grandmaster level performance in either chess, Go, or shogi (Japanese chess). The rapidly growing area of Generative Adversarial Networks (GANs) (Goodfellow et al., 2014) is based on the clever idea of pitting a generative neural network that produces objects against an adversary neural network that is tasked with discriminating natural objects from those produced by the generative network. One can imagine, say, a generative network receiving rewards based on how well its counterfeit bills fool a discriminator network, at the same time that the discriminator network receives rewards based on how accurately it accepts real money and rejects counterfeits. Over time, each network performs its task better. Systems of learning networks mutually testing and training each other in this way have produced systems capable of producing pictures of faces that are extremely realistic and novel, high-resolution images based on low-resolution inputs, photograph-like images of objects based on sketches or even text descriptions, realistically colorized black and white images, accurate high-resolution predictions for what the next frame of a movie will look like, and compelling artistic creations that combine one object's style with another object's content (Cresswell et al., 2018; Isola, Zhu, Zhou, & Efros, 2017; Kang, Gao, & Roth, 2019; Mazzone & Ahmed, 2019). In psychology, researchers have argued that human reasoning evolved in the adversarial social context of people trying to justify their beliefs and actions to others, and convince others to do their bidding through argumentation (Mercier & Sperber, 2019). GANs researchers have primarily been trying to solve specific engineering problems rather than create humanlike systems, and so it is striking that work on humans and computers has converged on adversarial processes for evolving agents to behave in increasingly sophisticated ways.

These are just two of many possible examples of the kind of inspiration that an agent-based perspective on systems fosters. For me, this perspective is a more compelling unifier for cognitive science than the "Mind as computer" slogan that Núñez et al. (2019)

describe as an early unifying theme, but which has also spurred considerable dissent of late, in part because of differences of opinion as to what computation involves. In any case, these examples demonstrate field-transcending connections made by researchers who are, for often times haphazard and serendipitous reasons, exposed to a theory in a very different field. The presence of interdisciplinary ventures like cognitive science can appreciably increase the odds of cross-fertilizations like these occurring (Repko, 2008).

4. Beyond the core

Cognitive science has a well-specified subject of inquiry (minds and intelligent systems), and this subject naturally requires the integration of disparate traditional fields. However, Núñez et al. (2019) make a convincing case that cognitive science does not possess many of the trappings of a traditional discipline. Different universities name and configure cognitive science in different ways, and have markedly different requirements for their students. The Cognitive Science Conference is not widely attended by researchers in many of the fields that cognitive scientists themselves see as highly relevant. The scarcity of anthropologists, philosophers, and computer scientists at the conference is conspicuous.

One sign that Núñez et al. interpret as cognitive science's failure to develop into a successful research program (Lakatos, 1978) is the increasing prevalence to referring to cognitive science*s* rather than cognitive science. Each variant reveals a value. The singular form stresses a common, unifying core (cf "The Commonwealth"), whereas the plural form reflects the strength that can be achieved by federating autonomous organizations into a larger unit (cf "The United States"). More generally, some of Núñez et al.'s analyses seem overly tied to the letter rather than spirit of cognitive science. For example, the graduate PhD degree in "Cognitive & Information Sciences" at University of California Merced is not included in their analysis of cognitive science graduate programs because it is not literally called "cognitive science," even though in terms of its content, it would be viewed as solidly within cognitive science by most self-described cognitive scientists. At the undergraduate level, Núñez et al. find that many degrees are based on a high percentage of coursework that does not explicitly mention "cognitive science" in its course titles (see their Fig. 4). This also is not problematic for cognitive science as a unified intellectual pursuit. The specific names given to courses are often based more on institutional history and chance than substantive content. A course on "Experiments and Models in Cognitive Science" and "Experiments and Models in Perception and Cognition" likely cover very similar material even though only the former counts as coursework in cognitive science by the authors' criterion. Few would deny that courses on "Classical Mechanics," "Electricity and Magnetism," "Quantum Mechanics," and "Thermodynamics and Statistical Mechanics" cohere to form the foundation of a physics curriculum even though none of these courses has "physics" in their titles. A deeper semantic and pragmatic analysis of intellectual institutions would be needed to determine whether there is a sufficient overlap in the training of cognitive scientists to have them coordinate their efforts on the project of understanding minds.

I suspect that there is, indeed, considerable overlap in the training of cognitive scientists nowadays. One can mention topics such a backpropagation, qualia, modularity, linguistic relativity, information theory, Turing machines, functionalism, and hippocampal place cells in cognitive science venues and have a good chance of being understood. Given the far-flung disciplinary origins of these concepts, it is impressive and remarkable that there is a sizeable community that understands them all at a reasonably deep level, and can even intelligibly make connections across the concepts.

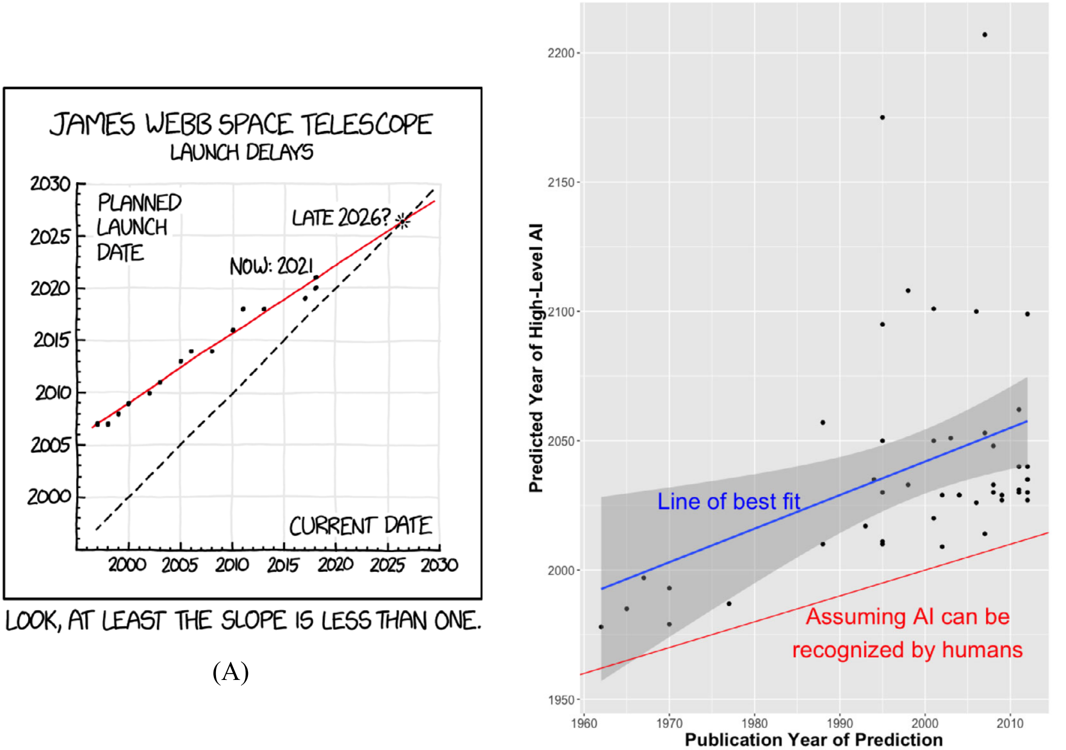
Beyond disputes about what counts as a training and research in cognitive science, a deeper way of engaging Núñez et al.'s institutional analysis is to question the extent to which a uniform disciplinary training across the world's cognitive science programs is a good thing. Cognitive science means different things in different places. Institutionally, cognitive science at UCSD has considerable strength in levels of analysis beyond the confines of individual people. University of California Irvine emphasizes computational models of cognition. MIT's program emphasizes neuroscience. Other institutions have emphases in vision, language, high-level cognition, or philosophy that distinguish them from other programs. In my opinion, this diversity of approaches should be celebrated. In much the same way that cross-cultural diversity confers resilience on humanity as a whole, so cross-institutional diversity in how cognitive science is approached confers resilience on the intellectual enterprise. In both cases, efforts to enforce greater uniformity across sites could well have the deleterious effect of reducing flexibility. When the problems facing a collective are difficult, it is often advantageous to increase the diversity of solutions by establishing local groups with greater within-group than between-group connectivity (Barkoczi & Galesic, 2016; Mason, Jones, & Goldstone, 2008; Page, 2017). And it is hard to imagine a tougher problem than an agent trying to understand its own mind using that same mind.

What other models for a successful cognitive science enterprise exist besides Lakatos's (1978) notion of intellectual troops rallying around and defending a hard, unchallenged core of assumptions and methods? One sharply contrasting model is one in which the attacks from within cognitive science on its own, original core have been healthy. In that spirit, the criticisms that Núñez et al. describe from embodied, embedded, extended, and enactive conceptions of mind (Colombo, Irvine, & Stapleton, 2019) toward the originally core tenet of "Mind as computation" have been healthy for putting cognitive science on a more robust foundation. According to the immunological "hygiene hypothesis," the increased prevalence of hay fever and asthma in many wealthy nations is due to the increased cleanliness of children's environments, which insufficiently challenge their developing immunological systems, causing their immune systems to respond inappropriately to innocuous substances (Yazdanbakhsh, Kremsner, & van Ree, 2002). The intellectual counterpart to this hypothesis asserts that research enterprises that are not adequately challenged fail to become robust. The attacks on the historic origins of cognitive science only serve to put cognitive science on a stronger foundation (Nietzche, 1889/1990), one which incorporates the relation between a mind and its environment, with environments prominently including tools constructed by the mind as well as other minds.

A second alternative to Lakatos's notion of a successful research program is to dispense with a common core and embrace pairwise cross-fertilization between disciplines. Cognitive science would have a valuable role to play in science even if it only served, say, to inspire machine learning researchers to consider promising approaches in human developmental neuroscience when constructing and analyzing their neural networks (Lillicrap & Körding, 2019), or psychologists to consider Bayesian models of probabilistic inference as plausible models of human judgment (Blanchard, Lombrozo, & Nichols, 2018). If we require consensus across the contributing disciplines of cognitive science about the meaning or proper application of a theory, then we will forfeit many opportunities for useful interdisciplinary connection and transfer. The highly pruned conceptual systems that survive a consensus alignment process are too meager to support fertile cross-pollination. By contrast, successful alignment and translation between pairs of conceptual systems within a larger set of systems is possible even when a single coherent alignment across all systems remains elusive (Goldstone & Rogosky, 2002). Under this view, cognitive science is the network community that emerges as a pattern of relatively dense local connections between diverse researchers studying minds and intelligent systems. One advantage of a network science approach to identifying a scientific community is that multiple, partially overlapping communities can be identified (Palla et al., 2005). The presence of established disciplines like computer science, anthropology, and psychology would not preclude the emergence of a network community of cognitive scientists.

A reasonable reply to these alternative conceptions of what it would mean for cognitive science to be intellectually vibrant is as follows: "Mathematics has a common core that its students are expected to know. Philosophy has its canon that includes Socrates, Descartes, and Husserl. Why shouldn't cognitive science have a common core?" Widely shared canons, practices, and norms do offer benefits. However, premature convergence on a shared core could have the unhealthy effect of reducing diversity, restricting otherwise rich explorations, and preventing useful partial convergences within subsets of fields bridged by cognitive science.

Premature convergence on a core would be particularly harmful for cognitive science because our enterprise is much closer to the "just getting started" than "moribund decline" stage of its lifespan. One striking illustration of this, shown in Fig. 1, comes from a compilation of AI experts' published predictions for when a high-level AI would be achieved (Armstrong, Sotala, & Ó hÉigeartaigh, 2014). For example, in 1988, Hans Moravec predicted human-like levels of intelligence in robots in the Year 2028. As disappointing as the delays in the James Webb Space telescope (Panel A) have been, at least the slope of the best fitting line relating the year of the forecast with the forecasted year is <1, suggesting a convergence around the Year 2026. By contrast, the slope of the best fitting line for a high-level AI is >1, completely failing to intersect with the $Y = X$ axis! Essentially all bets are off for when a human-like AI will be created. The gap between human and machine intelligence has not reduced over the years despite impressive achievements in AI such as playing chess and Go at grandmaster levels, autonomous



(A) (B)

Fig. 1. Panel (A) is from Randall Munroe (2018): https://xkcd.com/2014/. It shows a plot of forecasted launch dates for the long-delayed James Webb Space Telescope (from data available at https://en.wikipedia.org/wiki/James_Webb_Space_Telescope). The *X*-axis represents the date on which a forecast is made, and the *Y*-axis shows the forecasted year of completion on that date. Where the best-fitting line for the set of forecasts intersects the line $Y = X$ provides a prediction (the year 2026) for when the telescope will actually be launched. By contrast, expert predictions for when a high-level Artificial Intelligence will be built, assembled by Armstrong, Sotala, and Ó hÉigeartaigh (2014), are plotted in panel (B) with a best fitting blue line that has a slope >1, meaning that it never intersects with the red $Y = X$ line in the future.

driving, beating the best human players at question-answering games like Jeopardy, and recognizing natural speech and handwriting at highly effective levels. My own diagnosis for why the gap persists is that not enough attention has been paid to furnishing machines with (a) human-like constraints, both innate and self-acquired, (b) an ability to create rich and coherent models for generating predictions for situations not contained within their original training set, (c) mutually informing loops of generation and evaluation in which both processes are grounded in perception and action, and (d) a rich environment consisting crucially of a community of other agents similarly engaged in the process of making sense of the environment. Regardless of whether these specific criticisms are on target, it is clear that many exciting AI systems are being deployed to tackle important real-world tasks but also that the path towards creating humanlike AIs is still obscure. Our early explorations into creating working models of minds show promise, but we are also in need of fundamentally transformative approaches. In this context, fostering diversity of approaches, not rallying around a common core, is the strategy most likely to succeed for cognitive science.

It is not sensible to criticize the progress of cognitive science by claiming both that there is not yet a hard core to the cognitive science enterprise as one expects for traditional academic departments, and that cognitive science has failed to bridge the participating fields of philosophy, anthropology, linguistics, neuroscience, computer science, and psychology. If cognitive science were to become more like a conventional discipline, then it would likely also become more insular. Interdisciplinary journals tend to have broader, more diverse citations than do disciplinary journals (Leydesdorff & Rafols, 2011).

Personally, I am thankful that cognitive science is not a traditional discipline with a discipline's associated insularity and uncontested core. Humanity's collective pursuit of knowledge requires efforts on many points along a stable-to-fluid continuum. Like many cognitive scientists, I am frequently drawn more to the fluid, interdisciplinary pole. Yes, I would like cognitive scientists to be able to talk with each other deeply and efficiently, assuming, to some degree, a common knowledge base. But in the end, I find opportunities for discovering cross-disciplinary commonalities even more enticing. Maintaining a truly interdisciplinary status is inherently instable (van den Besselaar, 2018; Leydesdorff & Rafols, 2011), and the very real and immanent threat of cognitive science being completely absorbed into psychology is just one threat to cognitive science's fragile existence. Our venture must compete against other interdisciplinary enterprises like network science and complex systems. It must continue to be relevant to research outside of itself, and renew itself by bringing in different perspectives, even on matters that were once thought of as agreed upon. Still, there is an undeniable excitement that comes from being part of something that is still, and may always be, in the process of becoming (Barad, 2007).

Acknowledgments

The author thanks Bob French, Mirta Galesic, Dedre Gentner, Coty Gonzalez, David Kirsh, Paul Munro, Rafael Núñez, Henrik Olsson, Lael Schooler, Chris Schunn, Richard Shiffrin, Paul Smaldino, and Peter Todd for fun, informative, and helpful discussions.

References

Agmon, E., Gates, A. J., Churavy, V., & Beer, R. D. (2016). Exploring the space of viable configurations in a model of metabolism-boundary co-construction. *Artificial Life*, *22*, 153–171.

Armstrong, S., Sotala, K., & Ó hÉigeartaigh, S. S. (2014). The errors, insights and lessons of famous AI predictions – And what they mean for the future. *Journal of Experimental & Theoretical Artificial Intelligence*, *26*(3), 317–342.

Barad, K. (2007). *Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning*. Durham, NC: Duke University Press.

Barkoczi, D., & Galesic, M. (2016). Social learning strategies modify the effect of network structure on group performance. *Nature Communications*, *7*, 13109.

Bedau, M. A. (2003). Artificial life: organization, adaptation, and complex from the bottom up. *TRENDS in Cognitive Sciences*, *7*, 505–512.

Beller, S., & Bender, A. (2015). Exploring cognitive diversity: Anthropological perspectives on cognition. *Topics in Cognitive Science*, *7*, 548–551.

Blanchard, T., Lombrozo, T., & Nichols, S. (2018). Bayesian occam's razor is a razor of the people. *Cognitive Science*, *42*(4), 1345–1359.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced choice tasks. *Psychological Review*, *113*, 700–765.

Chiappori, P.-A., Levitt, S., & Groseclose, T. (2002). Testing mixed-strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer. *American Economic Review*, *92*(4), 1138.

Colombo, M., Irvine, E., & Stapleton, M. (2019). *Andy Clark and his critics*. Oxford, UK: Oxford University Press.

Cresswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. (2018). Generative adversarial networks: An overview. *IEEE Signal Processing*, *35*(1), 53–65.

Forstmann, B. U., Ratcliff, R., & Wagenmakers, E.-J. (2016). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annual Review of Psychology*, *67*, 641–666.

Gardner, H. (1987). *The mind's new science: A history of the cognitive revolution*. New York: Basic Books.

Gentner, D. (2010). Psychology in Cognitive Science: 1978–2038. *Topics in Cognitive Science*, *2*(3), 328–344.

Gold, I. J., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, *30*, 535–74.

Goldstone, R. L., & Leydesdorff, L. (2006). The import and export of Cognitive Science. *Cognitive Science*, *30*, 983–993.

Goldstone, R. L., & Rogosky, B. J. (2002). Using relations within conceptual systems to translate across conceptual systems. *Cognition*, *84*, 295–320.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. In *Advances in Neural Information Processing Systems (NIPS)* (pp. 2672–2680). Montreal, Canada.

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

Horowitz, J. M., & England, J. L. (2017). Spontaneous fine-tuning to environment in many-species chemical reaction networks. *Proceedings of the National Academy of Sciences*, *114*, 7565–7570.

Isola, P., Zhu, J., Zhou, T., & Efros, A. (2017). Image-to-image translation with conditional adversarial nets. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 1125–1134). Honolulu, HI.

Kang, Y., Gao, S., & Roth, R. (2019). Transferring multiscale map styles using generative adversarial networks. *International Journal of Cartography*, *5*(2–3), 115–141.

Lakatos, I. (1978). *The methodology of scientific research programmes* (vol. *I*). Cambridge, UK: Cambridge University Press.

Latty, T., & Beekman, M. (2011). Speed-accuracy trade-offs during foraging decisions in the acellular slime mould *Physarum polycephalum*. *Philosophical Transactions of the Royal Society. B*, *278*, 539–545.

Leydesdorff, L., & Goldstone, R. L. (2014). Interdisciplinarity at the journal and specialty level: The changing knowledge bases of the journal Cognitive Science. *Journal of the American Society for Information Science and Technology*, *65*, 164–177.

Leydesdorff, L., & Rafols, I. (2011). Indicators of the interdisciplinarity of journals: Diversity, centrality, and citations. *Journal of Informatics*, *5*, 87–100.

Lillicrap, T. P., & Körding, K. P. (2019). What does it mean to understand a neural network? *arXiv* 1907.06374v1.

Marshall, J. A. R., Bogacz, R., Dornhaus, A., Plaque, R., Kovacs, T., & Franks, N. R. (2011). On optimal decision-making in brains and social insect colonies. *Journal of the Royal Society Interface*, *6*, 1065–1074.

Mason, W. A., Jones, A., & Goldstone, R. L. (2008). Propagation of innovations in networked groups. *Journal of Experimental Psychology: General*, *137*, 422–433.

Maturana, H. R., & Varela, F. J. (1973/1980). *Autopoiesis and cognition: The realization of the living*. Dordrecht, Holland: D. Reidel.

Mazzone, M., & Ahmed, E. (2019). Art, creativity, and the potential of artificial intelligence. *Arts*, *8*, 26.

Mercier, H., & Sperber, D. (2019). *The enigma of reason*. Cambridge, MA: Harvard University Press.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York, NY: Henry Holt and Co. https://doi.org/10.1037/10039-000

Mitchell, M. (2009). *Complexity: A guided tour*. Oxford, UK: Oxford University Press.

Nietzche, F. (1889/1990). *Twilight of the Idols*. New York, NY: Penguin Classics.

Núñez, R., Allen, M., Gao, R., Miller Rigoli, C., Relaford-Doyle, J., & Semenuks, A. (2019). What happened to cognitive science? *Nature Human Behavior*, *3*, 782–791. https://doi.org/10.1038/s41562-019-0626-2

Page, S. E. (2017). *The diversity bonus: How great teams pay off in the knowledge economy*. Princeton, NJ: Princeton University Press.

Page, S. E., & Miller, J. (2007). *Complex adaptive systems: An introduction of computational models of social life*. Princeton, NJ: Princeton University Press.

Palla, G., Derényi, I., Farkas, I., & Vicsek, T. (2005). Uncovering the overlapping community structure of complex networks in nature and society. *Nature*, *435*(7043), 814–818.

Pinker, S. (1997). *How the mind works*. New York, NY: W. W. Norton & Company.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.

Repko, A. F. (2008). *Interdisciplinary research: Process and theory*. Los Angeles, CA: Sage.

Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, *44*, 206–226.

Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T., Simonyan, K., & Hassabis, D. (2018). A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. *Science*, *362*(6419), 1140–1144.

Simon, H. A. (1969). *The sciences of the artificial*. Cambridge, MA: MIT Press.

Thurner, S., Klimek, P., & Hanel, R. (2018). *Introduction to the theory of complex systems*. Oxford, UK: Oxford University Press.

van den Besselaar, P. (2018). Interdisciplinary and disciplinary identities: towards a theory of forms of knowledge change. *Preprint at bioRxiv*, https://doi.org/10.1101/603449.

von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Vugteveen, P., Lenders, R., & Van den Besselaar, P. (2014). The dynamics of interdisciplinary research fields: the case of river studies. *Scientometrics*, *100*, 73–96.

Yazdanbakhsh, M., Kremsner, P. G., & van Ree, R. (2002). Allergy, parasites, and the hygiene hypothesis. *Science*, *19*, 490–494.