TOPICS
TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science 11 (2019) 918–927

ISSN:1756-8765 online
DOI: 10.1111/tops.12468

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

# Expanding and Repositioning Cognitive Science

Paul S. Rosenbloom,[a] Kenneth D. Forbus[b]

[a]*Department of Computer Science & Institute for Creative Technologies, University of Southern California*
[b]*Computer Science Department, Northwestern University*

Received 4 August 2019; received in revised form 16 September 2019; accepted 16 September 2019

**Abstract**

Cognitive science has converged in many ways with cognitive psychology, but while also maintaining a distinctive interdisciplinary nature. Here we further characterize this existing state of the field before proposing how it might be reconceptualized toward a broader and more distinct, and thus more stable, position in the realm of sciences.

*Keywords:* Cognitive science; Computation; Artificial intelligence; Animal cognition; Methodology

## 1. Introduction

As Gentner (2010) and Núñez et al. (2019) have observed, cognitive science has become dominated by cognitive psychology, to the detriment of the original goals of the field. But what were those goals? Núñez et al. quote Gardner (1987) to the effect that cognitive science should have led to a dissolution of the borders between it and the fields out of which it was originally to be composed, and then expresses concern that this has not happened. In addition to this strong hypothesis about what cognitive science should be, Núñez et al. also quote Gardner’s weak hypothesis about the structure of the field, which just says it involves cooperation among the participating disciplines.

Correspondence should be sent to Paul S. Rosenbloom, Department of Computer Science & Institute for Creative Technologies, University of Southern California, 10215 Waterfront Drive, Playa Vista, CA 90094. E-mail: rosenbloom@usc.edu

It is our sense that cognitive science today falls between these two hypotheses, embodying a more moderate outcome under which the field maintains a distinct identity but without taking over the constituent fields. Section 2 characterizes further this current state, including how cognitive science today is necessarily most strongly bonded to cognitive psychology due to their shared goal of understanding human cognition. Shifting this center of gravity requires expanding on this goal, a topic that is the focus of Section 3. Section 4 then goes beyond this to discuss expanding the set of methods employed toward these goals. Although such methodological diversity could be considered even without expanding the goals, the present goal alignment has, as a side effect, largely aligned the methods as well. As summarized in Section 5, combining these two forms of expansion has the potential to yield a broader and more unique field with a distinct shift in its center of gravity.

## 2. Cognitive science today

Núñez et al. quote Simon and Kaplan (1993) to the effect that cognitive science is "the study of intelligence and intelligent systems, with particular reference to intelligent behaviour as computation." Although it is not completely clear from this quote whether Simon and Kaplan intended to include just human intelligence in this definition, a later quote in the same paper does clarify this to mean "in men (and animals), in computers, and in the abstract." Despite this, it is clear that cognitive science today focuses almost exclusively on intelligence in humans, with the other organisms and systems only relevant to the extent that they contribute to understanding human intelligence. The Cognitive Science Society itself says that it "brings together researchers from around the world who hold a common goal: understanding the nature of the human mind" (https://cognitivesciencesociety.org/about/, September 6, 2019).

The role of computation in the production of intelligent behavior might have provided an additional differentiator between cognitive science and cognitive psychology. However, the use of computation has simultaneously become more common within cognitive psychology, possibly due to the influence of cognitive science, and less central to cognitive science, as witnessed by the quote above from the Society that omits it from the "common goal" of cognitive science. Thus, very little now distinguishes the goals of the two disciplines, and so it should not be at all surprising that cognitive science has become closely aligned with cognitive psychology, while maintaining at best an interdisciplinary relationship between it and the other related disciplines that all maintain different goals.

We are also therefore not surprised at the lack of cognitive science departments, since most interdisciplinary programs are not departments. Training every graduate student in every discipline to become an expert in them all seems a fool's errand, due to the time it would take. Training PhDs in a particular discipline's methods is known to be doable within the normal span of a graduate program. Thus, the typical approach in cognitive science is to train students to be world-class in at least one discipline's methods, and have a working knowledge of others, so that they can combine them as appropriate and collaborate effectively across traditional disciplines. Science is more of a team sport these days and ensuring that students can pursue research productively with scholars of varied backgrounds is something that cognitive science training today does reasonably well.

This approach thus goes modestly beyond Gardner's weak hypothesis in proactively training new generations of students for working across multiple disciplines. It furthermore goes beyond this weak hypothesis in yielding a theoretical language for cognitive science that is a pidgin; that is, a rather crude mixture of multiple source languages. In cognitive science papers, one sees amalgams of terms and ideas from multiple disciplines. Based on observing graduate students and young researchers in cognitive science, we believe that this language is on its way to becoming a creole; that is, a stable language that is spoken natively by subsequent generations. Cognitive science is still young and is tackling one of the hardest problems on which humanity has ever worked, so this pace should be unsurprising. Still, it is further evidence of cognitive science already becoming established as an interdisciplinary field that goes beyond mere cooperation.

In summary, cognitive science today is thus largely aligned with cognitive psychology due to how their goals align. However, it remains an interdisciplinary field that goes beyond simple cooperation among the constituent disciplines in both its educational structure and the language it uses.

## 3. Expanding the goals

For cognitive science to be something more than just an explicitly interdisciplinary variant of cognitive psychology, its goals must be distinct in some significant way. Simon and Kaplan's earlier discussion of the nature of cognitive science turns out to supply two such possibilities, one that we will firmly reject here and the other that we will strongly advocate.

The first concerns the role of computation in cognitive science. Could reestablishing it to its earlier definitional role in the field provide the necessary distinction between it and cognitive psychology? Pragmatically this would be difficult due to the movement that has already been mentioned of cognitive psychology in this direction. However, more fundamentally, a field is best defined in terms of the phenomena it studies, with neither methods nor hypotheses properly part of this. Instead, methods and hypotheses generally evolve as the understanding of the phenomena improves. Methods go out of style, but fields do not, or at least should not: Physics and biology, for example, are still being studied, even though the hypotheses have evolved, and the methods used to study them have changed dramatically. Likewise, hypotheses—such as that cognition is computation—may ultimately pan out or not, but their status should remain a scientific question rather than a definitional one for the field.

As witnessed by the quote from Simon and Kaplan, the power of computation historically was one of the main motivations for founding cognitive science, and it does provide both relevant methods and hypotheses. With respect to methods, computation provides a powerful language for expressing cognitive theories, creating cognitive models, and experimenting with cognitive models to determine their implications. With respect to hypotheses, whether computation is the key to producing intelligence is central. However, neither truly justifies the notion of computation as being definitional for the field. Several threads of cognitive science—such as situated cognition, distributed cognition, and embodied cognition—have, in fact, defined themselves partly in terms of opposition to what they viewed as computation (even given work on how computation can bear an interesting relationship to such ideas, as their adaptation by AI researchers demonstrates).

In this same issue, Cooper (2019) mentions the notion of "research traditions," as articulated by Laudan (1977). The notion of computation as being central to intelligence and intelligent behavior may ultimately be best considered as such a tradition, as with Laudan's examples of the role of quantum theory in physics and of evolutionary theory in biology. In this manner, computation would remain a fully acknowledged major force in cognitive science, but one whose status could be questioned within the field itself. At the same time, it would also clearly not be definitional for the field.

The second possibility from Simon and Kaplan is to go with "intelligence and intelligent behavior" "in men (and animals), in computers, and in the abstract," although with the substitution of "humans" for the archaically gendered term of "men." In their article, Simon and Kaplan mentioned that "no really satisfactory intensional definition of intelligence has been proposed," and unfortunately there is still no generally accepted definition, despite a series of attempts that have yielded more sophisticated approaches over time. We will not attempt to make further progress on this question here; however, we do expect to see continued improvements in such attempts as the field matures in its overall understanding of what constitutes intelligence. In effect, the development of the definition is anticipated to proceed hand-in-hand with the understanding of it as a matter for study, rather than for the former to precede the latter. This ultimately is not too different from how biology has operated with respect to defining the concept of life; as life becomes better understood, its definition becomes clearer.

Starting with animals, there is an increasing, and increasingly useful, trend of comparing cognition across species, for example, examining the capabilities of human and nonhuman primates (Christie et al., 2016; Seed & Tomasello, 2010), corvids (Emery & Clayton, 2004; Logan et al., 2014), and dogs (Hare & Tomasello, 2005), to better understand the nature of cognition. We believe that this should become—and perhaps should always have been—an essential part of cognitive science. But, as with Simon and Kaplan, we also propose to go further, to include both the actuality and potentiality of intelligence in constructed entities; and, in fact, intelligence in the abstract. Insights from studying particular organisms and entities can help in understanding other organisms and entities, with an understanding of the overall space of possibilities and how it is structured possibly contributing even more.

Artificial intelligence becomes as central to such a field as does cognitive psychology, with other fields that study animals or optimal approaches to problems—such as operations research—also becoming relevant. This ultimately articulates a problem—of understanding intelligence in all existing and possible forms plus the structured space of intelligences that they comprise—and an associated body of phenomena that could define the next best grand challenge, after the understanding of human intelligence.

In summary, the proposal here is to solidify the omission of the notion of computation as definitional for the field of cognitive science—reformulating it instead as an important research tradition—and re-expanding the goals of the field to include the structured space of all existing and possible instantiations of intelligence and intelligent behavior in humans, animals, machines, and in the abstract. Some of these forms of intelligence might be quite trivial, as in snails or thermostats, but they would still be part of the full space of possibilities. Some of them also clearly implicate thought in constructed systems, with cognitive science thus becoming a science of both the natural and the artificial (Simon, 1969).

## 4. Expanding the methods

Cognitive psychology predominantly focuses on experimentation with humans, and cognitive science with a similar goal has largely followed its lead on this as well. Expanding this to experimentation with the full breadth of intelligent systems, both natural and artificial, is thus one particularly straightforward way for cognitive science to go beyond the methods of cognitive psychology.

Theory and models can also be developed in cognitive psychology, but they are most often tied tightly to the experimental data on which they bear. Although this does provide an admirably cautious way of proceeding, and caution is terribly important in science when drawing conclusions upon which others will depend, theory must at times have the freedom to diverge from what is currently known experimentally to be able to reach its full potential. It must be able to consider what is not yet known, as well as to abstract over what is currently known, in creating broad theoretical edifices—such as logics, learning paradigms, cognitive architectures (Kotseruba & Tsotsos, 2018) or the Common Model of Cognition (Laird, Lebiere, & Rosenbloom, 2017)—that can provide essential overall guidance to the field. To cite just one recent example, the Common Model of Cognition—an attempt at developing a community consensus as to the structures and processes necessary to yield a human-like mind—has been used to guide the development of a model of functional connectivity between brain regions that, when evaluated against fMRI data, outperformed traditional approaches (Stocco et al., 2018).

Theory must be held as important as experimentation across the sciences—with physics providing an obvious exemplar of a discipline where it is held as such—and must also have the ability to proceed as necessary independently of experiments, just as experiments at times must be able to explore new phenomena even prior to any existing relevant theory. Ultimately, theory and experiments must fully co-articulate, although forcing them to always do so, irrespective of the current stage of understanding, is not necessarily the best research strategy. AI can, for example, bring computational complexity arguments, and sometimes even representational constraints, to bear on understanding cognition prior to the availability of relevant human data. For instance, predictions from qualitative process theory concerning directionality in human causal reasoning about quantities have held up quite well across many domains (Forbus, 2019).

Should theory take on a more prominent and independent position within cognitive science, both philosophy and AI should also take on more prominent roles within the field. Cooper (2019) points out that philosophy actually plays more of a role in cognitive science than is apparent from the analysis by Núñez et al., but the proposal here suggests a potentially even larger role.

Applications, or even abstract task-domain analyses that fall short of producing working systems, have played even less of a role than theory methodologically in cognitive psychology, or in a cognitive science that has become difficult to distinguish from cognitive psychology. Simon (1969), with his ant analogy, early on pointed out how much of the complexity of intelligent behavior can be due to the complexity of the environment rather than the complexity of the organism. More recent work on rational analysis (Anderson, 1991), computational rationality (Lewis, Howes, & Singh, 2014), and Bayesian modeling (Griffiths et al., 2010) has taken this fully to heart, in exploring how human behavior can be considered rational given its goals and environment (and, in some cases, its embodiment). Thus, understanding of intelligent behavior by necessity requires understanding both task domains and specific tasks.

What role should applications and more abstract task analyses play? Many modern researchers do not consider them to ever have a proper role in science, relegating them instead to engineering or its ilk. However, compelling arguments can be made for the importance of both use-inspired research in what is termed Pasteur's Quadrant (Stokes, 1997) and how great scientific domains properly include both understanding and shaping, whether in the form of physical science + engineering, or life science + medicine, or social science + business (or law) (Rosenbloom, 2012). Applications can not only produce useful results that go beyond the realm of science, but also within science they can help evaluate and provide feedback on how well theories work and suggest areas where more understanding is required. This is another general area in which AI can contribute in significant ways to cognitive science. For example, the notion that deep learning may be more useful to cognitive science in providing abstract analyses of human tasks rather than actual models of human behavior was heard from multiple speakers at the most recent conference of the Cognitive Science Society.

If the methods of cognitive science are extended to include a more symmetric approach that includes not only theory as a central component on its own that is not simply ancillary to experiments, but also applications—along with more abstract analyses of task domain—it must be acknowledged that additional forms of evaluation will be needed for these methods. This is a tricky topic in general, as such methods often do not support evaluations that are as strong—in terms of how much confidence they can provide in the validity of their results—yet it is critical that such key topics not be relegated to obscurity just because they are more difficult to evaluate; otherwise, every method not involving mathematical proof could ultimately get pushed aside, as not even careful experimentation is as strong a method as mathematical proof.

Understanding the necessary range of methods and what can legitimately be concluded based on the strength of their evaluations would need to become a necessary part of the field's overall expertise and curriculum. For theories, general approaches to their evaluation can be based on, for example: fits to existing data; the ability to predict new data; the potential for falsifiability, even when such data is not yet available; simplicity, as in Occam's razor; beauty/elegance, as is common in physics (e.g., Wilczek, 2015) although controversial there as well (Hossenfelder, 2018), and as has recently been proposed as one of the driving desiderata for the Sigma cognitive architecture (Rosenbloom, Demski, & Ustun, 2016); explanatory reach (Deutsch, 2011); sufficiency/ability, in terms of whether the theory actually supports task performance (Cassimatis, Bello, & Langley, 2008); and community consensus, as is sought with the Common Model of Cognition.

Not all methods can be used on all theories, and different methods clearly vary in what they guarantee about the theories to which they can be applied, but the key point here is that there is a variety of types of evaluations that can be applied to yield insight into them, even before the availability of appropriate human data. To take one specific example mentioned briefly above that is particularly relevant to cognitive science, applications can support an important form of sufficiency/ability evaluation for theories in cognitive science, by helping to determine whether they can actually produce intelligent behavior versus just predicting aspects of that behavior. AI systems are held to a performance standard: If a theory concerns how to do a task, a machine must be able to use an implementation of that theory (i.e., its processes and representations) to actually do that task; and, in the process, much can be learned about both the theory and the task. For example, as qualitative reasoning systems have become used in scientific and engineering practice, the kinds of representations and reasoning required have become greatly clarified (Forbus, 2019).

Many computational models in cognitive science fail by this standard: They might, for example, predict the reaction time of someone solving a problem, but not what the solution actually is. Consider the classical example of Fitts' Law (1954), which accurately predicts the time it takes for a human to move a pointer to an object as a function of the size of the object and its distance, but which has no way of actually moving a pointer to any object. Including such a sufficiency/ability measure for cognitive theories adds an important form of evaluation that is typically missing in cognitive psychology, but which should be much more central in cognitive science.

In other words, our proposal is to expand the methods used in cognitive science to include: (a) theory in its own right, and not just as something that is ancillary to experiments; and (b) applications and abstract task analyses. As the previous proposal concerning goal expansion harks back to early arguments by Kaplan and Simon, this proposal concerning method expansion harks back to even earlier arguments by Newell (1973), who advocated for more of a theoretical focus in cognitive psychology, particularly focused on models of the overall control structure of the mind, and for the study of complete, complex tasks (i.e., applications).

Given the discussion in the previous section, such expansions should not impact the definition of the field, but they should impact how the phenomena that do define the field may be studied. Such expansions do also need to be accompanied by additional work on appropriate forms of evaluation, for theoretical work that is to go beyond simply the degree of match or the ability to predict experimental results, and for application work that is to go beyond providing useful systems to illuminate the fundamental nature of intelligence.

## 5. Summary, proposals, and recommendations

In summary, we have characterized a moderate hypothesis concerning the current structure of the field of cognitive science and its relationship to its constituent fields. We have also made three proposals for the future of the field. The first is a radically expanded goal that spans understanding a broad range of phenomena concerned with intelligence and intelligent behavior across both natural organisms and artificial entities. The second is a field of cognitive science in which computation amounts to a research tradition that yields important methods and hypotheses but is not part of the definition of the field. The third is an expansion of the methods fully accepted by the field to admit more research on theories, task domains, and applications that are not tied to the details of human laboratory data. One clear consequence of these proposals is a greatly increased role for artificial intelligence in cognitive science, possibly to a position that is symmetric to the current role of cognitive psychology. Other fields, from animal psychology to philosophy, would likely also take on increased roles, but we are personally less well qualified to comment on these.

On a smaller scale, but with more immediate and concrete impact, we recommend including a broader range of AI members on the editorial boards of *Cognitive Science* and *Topics in Cognitive Science*. This will be necessary to support the longer-term manners in which AI should play a more central role in cognitive science, but also in the shorter term it should help start a shift towards welcoming a broader range of work. Something similar also seems appropriate for the annual conference, where papers that are not tied to specific human-subjects data appear to be routinely rejected. The discussion here also makes a general case for more participation by philosophers, and likely this argument could and should be extended to other disciplines as well, but we are less in a position to be concrete with respect to such recommendations.

## Acknowledgments

This effort has been sponsored by the U.S. Army, the Office of Naval Research, and the Air Force Office of Scientific Research. Statements and opinions expressed do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

## References

Anderson, J. R. (1991). Is human cognition adaptive? *Behavioral and Brain Sciences*, *14*, 471–517.

Cassimatis, N., Bello, P., & Langley, P. (2008). Ability, breadth, and parsimony in computational models of higher-order cognition. *Cognitive Science*, *32*, 1304–1322.

Christie, S., Gentner, D., Call, J., & Haun, D. B. M. (2016). Sensitivity to relational similarity and object similarity in apes and children. *Current Biology*, *26*(4), 531–535.

Cooper, R. P. (2019). The multidisciplinary flux of cognitive science. *Topics in Cognitive Science*, *11*(4), 869–879.

Deutsch, D. (2011). *The beginning of infinity: Explanations that transform the world*. London: Penguin Books.

Emery, N., & Clayton, N. (2004). The mentality of crows: Convergent evolution of intelligence in corvids and apes. *Science*, *306*, 1903–1907.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *47*, 381–391.

Forbus, K. (2019). *Qualitative representations: How people reason and learn about the continuous world*. Cambridge, MA: MIT Press.

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 to 2038. *Topics in Cognitive Science*, *2*, 328–344.

Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010). Probabilistic models of cognition: exploring representations and inductive biases. *Trends in Cognitive Sciences*, *14*, 357–364.

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

Hossenfelder, S. (2018). *Lost in math: How beauty leads physics astray*. New York: Basic Books.

Kotseruba, I., & Tsotsos, J. K. (2018). 40 years of cognitive architectures: Core cognitive abilities and practical applications. *Artificial Intelligence Review*, https://doi.org/10.1007/s10462-018-9646-y

Laird, J. E., Lebiere, C., & Rosenbloom, P. S. (2017). A standard model of the mind: Toward a common computational framework across Artificial Intelligence, Cognitive Science, Neuroscience, and Robotics. *AI Magazine*, *38*, 13–26.

Laudan, L. (1977). *Progress and its problems: Towards a theory of scientific growth*. Berkeley: University of California Press.

Lewis, R. L., Howes, A., & Singh, S. (2014). Computational rationality: Linking mechanism and behavior through bounded utility maximization. *Topics in Cognitive Science*, *6*, 279–311.

Logan, C. J., Jelbert, S. A., Breen, A. J., Gray, R. D., & Taylor, A. H. (2014). Modifications to the Aesop's fable paradigm change new caledonian crow performances. *PLoS ONE*, *9*, e103049.

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

Núñez, R., Allen, M., Gao, R., Rigoli, C. M., Relaford-Doyle, J., & Semenuks, A. (2019). What happened to cognitive science? *Nature Human Behaviour*, *3*(8), 782–791.

Rosenbloom, P. S. (2012). *On computing: The fourth great scientific domain*. Cambridge, MA: MIT Press.

Rosenbloom, P. S., Demski, A., & Ustun, V. (2016). The Sigma cognitive architecture and system: Towards functionally elegant grand unification. *Journal of Artificial General Intelligence*, *7*, 1–103.

Seed, A., & Tomasello, M. (2010). Primate cognition. *Topics in Cognitive Science*, *2*, 407–419.

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

Simon, H. A., & Kaplan, C. A. (1993). Foundations of cognitive science. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 154–170). Cambridge, MA: MIT Press.

Stocco, A., Laird, J., Lebiere, C., & Rosenbloom, P. (2018). Empirical evidence from neuroimaging data for a standard model of the mind. In T. T. Rogers, M. Rau, X. Zhu, & C. W. Kalish (Eds.), *Proceedings of* *the 40th Annual Meeting of the Cognitive Science Society* (pp. 1094–1099). Austin, TX: Cognitive Science Society.

Stokes, D. E. (1997). *Pasteur's quadrant: Basic science and technological innovation*. Washington, DC: Brookings Institution Press.

Wilczek, F. (2015). *A beautiful question: Finding nature's deep design*. New York: Penguin Books.