Title
The Narrow Conception of Computational Psychology

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 39(0)

Publication Date
2017

Peer reviewed

The Narrow Conception of Computational Psychology

Abstract

One particularly successful approach to modeling within cognitive science is computational psychology. Computational psychology explores psychological processes by building and testing computational models with human data. In this paper, it is argued that a specific approach to understanding computation, what is called the 'narrow conception', has problematically limited the kinds of models, theories, and explanations that are offered within computational psychology. After raising two problems for the narrow conception, an alternative, 'wide approach' to computational psychology is proposed.

Keywords: narrow conception, individualism, computation, psychology, explanation

Introduction

Cognitive science has gained a good deal of theoretical and methodological impetus from thinking about how psychological processes can be described, studied, and simulated using different types of models. One particularly successful approach to modeling is computational cognitive modeling or, more simply, computational psychology. Computational psychology explores psychological processes by building and testing computational models with human data (Sun, 2008).

In this paper, it is argued that a specific approach to understanding computation, what is dubbed the 'narrow conception', has problematically limited the kinds of models, theories, and explanations that are sometimes offered within computational psychology.

The impetus for the current study arises from a growing debate around the role, nature and status of computation within psychological investigations. Several authors have begun to re-examine what computationalism stands to offer the cognate disciplines (Piccinini, 2015; Milkowski, 2015). The current discussion stands to contribute to this growing trend by exploring and examining one important assumption that underwrites a notable swath of research within computational psychology. The goal is to show that computational psychology has overlooked an important constraining assumption.

Computational Psychology

For many, computational theory provides a theoretically flexible and expressively powerful tool for exploring cognition (Anderson 1983; Pylyshyn, 1984; Newell, 1990; Anderson & Lebiere, 1998, 2003). The computational approach allows researchers to construct detailed accounts of the mechanisms, structures, and processes that underwrite cognition. In testing and extending the theories of other domains, such as cognitive psychology and artificial intelligence, computational investigations offer a functionally viable yet mathematically rigorous way of exploring cognitive or psychological processes.

A good deal of the explanatory value of computational psychology lies not only in the ability to produce computer simulations, but also in using those simulations to make predictions about human data. By matching the 'fit' of human data with computer simulations, researchers establish systematic relationships between computational models and psychological processes, which can reveal the underlying structure and form of cognitive functionalities (Sun & Ling, 1998).

Consider three illustrative examples of computational psychology in action. First, consider Shiffrin and Steyvers' (1997) REM model of episodic memory. Shiffrin and Steyvers' model is one instance of a class of abstract, computational models that attempt to explain recognition judgments. These models employ a 'global matching' procedure. The global matching procedure produces a familiarity signal that indicates whether or not an item has been previously presented to the model – a test cue, for example, that matches two features of one item will yield a higher familiarity judgment than a test cue that matches one feature of each of the two items.

Shiffrin and Steyvers' model puts a Bayesian twist on the global matching procedure. The REM model calculates the likelihood of whether a cue item matches or corresponds to particular stored memory traces by assigning values to each of the stored items. When the model is tested to see if it can identify whether cue items are new or old, the cues are compared with each trace item in memory such that the model calculates the likelihood of the retrieval cue and the trace item matching. Recognition judgment is explained in terms of a probabilistic familiarity process operating within memory.

According to Shiffrin and Steyvers, the REM model accounts for a number of distinct memory effects. One example is the word frequency mirror effect. The word frequency mirror effect says that subjects often make more false alarms on high-frequency lures (foils) versus low-frequency lures and more correct "old" responses to low-frequency targets versus high-frequency target when making recognition judgments (Glanzer et al., 1993). The REM is able to accommodate the word frequency effect in virtue of the fact that low-frequency words have more unusual features than high-frequency words (e.g., more syllables). The REM model is able to use a slightly lower value when generating low frequency lures items during matching, which results in these items having slightly higher feature values.

The relevance of the REM model is that in measuring the fit of the model to behavioral data and by adopting a Bayesian approach to the global matching procedure, the REM model focuses on both the essential interplay between modeling and experimental data and formalizing cognitive processes in a computationally rigorous way.

Next, consider Dienes' (1992) connectionist model of implicit language learning. Dienes' model attempts to computationally instantiate how language users come to implicitly understand artificial grammars using artificial neural networks. In particular, Dienes' model uses a feed-forward autoassociator network.

In a feed-forward autoassociator network – which is a version of the more standard multilayer network – activation passes through the connection weights of the network just once to produce the output activation. The feed-forward autoassociator contrasts with recurrent autoassociator networks, in which the output activation arrives back at each node and is passed through the weights again until a stable state is reached.

In terms of network training, Dienes' model was presented with the same learning material as human subjects, which included arbitrary strings of letters, such as MTTTTV or MTTVT. These features of strings were represented as activations to the network's input layer. Depending on whether the feature was present or absent, the unit coding the feature would have an activation of either 1 or 0. Once the network learned the arbitrary training strings, similar to experimental tests, the model was then made to make grammaticality judgments on new strings of letters. The goal was to see if the network had learnt the underlying grammatical principles that implicitly structured the arbitrary strings being presented.

When Dienes' tested the model, it was found the network was able to distinguish grammatical versus ungrammatical strings. The network was able to reproduce the training strings by adding or subtracting strings from an exemplar case. The model predicted each feature of a string based on some set of the remaining features from an exemplar. When Dienes' compared network results to that of human subjects, it was found that the network could classify test strings as well as people could. The network tended to reproduce grammatical test strings more faithfully than non-grammatical test strings.

Dienes' connectionist model stands as a further interesting example of computational research, as it provides a computational account of implicit artificial grammar learning that measures the fit of the model with behavioral data. By investigating how artificial neural networks handle artificial grammar tasks, Dienes' attempts to undercover the computational processes and representations underlying implicit language learning.

A third example comes from Osherson et al.'s (1990) declarative model of inductive reasoning. Osherson et al.'s (1990) model attempts to investigate the computational underpinnings of 'inductive' reasoning – inductive reasoning is the process by which premises are thought to lend non-conclusive support to the truth of specific conclusions.

In Osherson et al.' model, inductive reasoning is explained in terms of the assessment of propositional statements according to the similarity between premise and conclusion categories For example, consider two inferential chains: (i) Mice have property X/All mammals have property X and (ii) Horses have property X/All mammals have property X. The category of 'mammal' in the conclusion covers both mice and horses. For Osherson et al., understanding how humans are able to make inferences about mice and horses depends on understanding how structural relationships between different categories are established – for example, understanding that mice and horses are instances of the subordinate category mammal.

Two features allow Osherson et al.'s model to make sense of cases such as the above. The first is that the model assesses the similarity between premise categories and conclusion categories. The second is that the model measures how well the premise categories covers the superordinate category. Coverage between premise and conclusion categories is assessed in terms of the average similarity of the premise category to members of the superordinate category. For instance, to the extent that horses are more typical mammals than mice, and therefore more similar to other kinds of mammals, (ii) will have greater coverage than (i).

Osherson et al.'s model is interesting because it addresses a number of empirical phenomena. One example is similarity effects. Similarity effects occur when people make inferences based on the perceived similarity between items in different inferential chains. Osherson et al. (2008) found, for example, that when people were given a choice between two syllogistic arguments about 95% chose the argument that they perceived to contain the greater similarity between premise and conclusion categories, e.g., sparrows to robins and blue jays versus geese to robins and blue jays. Osherson et al.'s model was able to accommodate such cases by assessing the relationship between the subordinate and premise categories.

Similar to the previous models, Osherson et al.'s declarative model is an illustrative example of computational psychology, because it is not only informed by and tested against empirical data, but it also attempts to identify the computational procedures and properties underlying complex cognitive processes, such as inductive reasoning.

The point of the previous survey is that each of the three models provides a paradigmatic example of computational psychology. Each model attempts to undercover the computational underpinnings of various cognitive processes via the construction and testing of computer models with human data. These models help to tease out the underwriting assumptions within computational research.

The Narrow Conception

With the domain of analysis laid out, the task now is to examine one approach to understanding computation that underlies a good deal of the research within computational psychology, what is labeled the 'narrow conception'.

In order to get a better handle on the narrow conception, consider what Segal (1991) says about computational cognitive systems, he writes: "It seems likely that whole subjects (or whole brains) make up large, integrated, computational systems…the whole subject is the largest acceptable candidate for the supervenience base because it is the largest integrated system available" (p.492). For Segal, the individual or whole subject (which is plausibly identical to the whole brain) is the largest unit available for computational, psychological investigation. Newell et al. (1989) offer a similar view, writing: "Symbol systems are an interior milieu, protected from the external world, in which information processing in the service of the organism can proceed" (1989, p.107). Here, again, computational systems are limited to the boundary of the individual.

Consider Fodor (1983) next: "Mechanisms of transduction are thus contrasted with computational mechanisms: whereas the latter may perform quite complicated, inference-like transformations – the former are supposed – at least ideally – to preserve the information content of their input" (1983, p. 41). Fodor's contrast between sensory transducers and computational mechanisms is indicative of where he thinks computational systems are located. Computational systems are sandwiched between transducers and motor outputs. Finally, consider what Egan (2000) says on the matter: "A computational theory prescinds from the actual environment because it aims to provide an abstract, and hence completely general, description of a mechanism that affords a basis for predicting and explaining its behaviour" (p.191). Only by abstracting away from the embedding environment and focusing on the individual can one begin to provide successful computational analyses. Once again, the outer limit of formal analysis for computational systems is the individual.

Common to each of these views is the idea that the individual or some sub-module, conceived of in terms of the primary unit of action, constitutes the largest organizational system amendable to computational description (i.e. computational modeling). The individual marks the conceptual boundary for computational, psychological investigations. Here is one way the view might be formulated:

> THE NARROW CONCEPTION: Computational cognitive systems are, and should be studied as if they were, located entirely within the individual or some sub-module.

Something in the spirit of this claim seems to have operated implicitly within a good swath of computational psychology. The narrow conception, if true, represents a principled claim about where and how computational cognitive systems should be studied. It constitutes a plausible and substantive proposal for computational psychology.

Consider the methodological implications of the narrow conception. If computational systems are wholly interior to the individual, then computational modeling should have as its target only those systems and processes that are individual-centered. As Segal diagnosis the situation: "Whole subjects plus embedding environments do not make up integrated, computational systems" (1991, p.492). The embedding environment plus individual will always fail to be adequate for computational analysis. Only the individual or some sub-system will be sufficient for computational modeling.

One motivation for adopting the narrow conception is that it provides a powerful way of explaining the causal powers of cognition. If cognitive systems are computational systems and computational systems are located within the individual, then identifying the causal properties and powers of computational systems provides insight into causal power of cognitive processes and abilities. Memory effects, such as primacy and recency affects, for example, will be best explained by focusing on the computational search strategies used by individuals during various tasks (e.g., exhaustive versus terminal search) (Sternberg, 1969). Only by identifying the distinct functional and causal properties intrinsic to the individual are rigorous computational, psychological explanations provided.

What is interesting about the narrow conception, besides its relatively straightforward nature, is that it is plausibly supported by and conforms to a good deal of research within computational psychology. This is why authors such as Segal claim that it is "likely" that the whole subject is the largest unit of analysis. The narrow approach is an empirical wager on how computational cognitive systems are distributed in nature.

Return to the three previous models to see why. First, consider how Shiffrin and Stevyer describe their model: "This cued recall model is meant to illustrate one plausible way in which retrieval from episodic images and retrieval from lexical/semantic images could work hand in hand to allow recall to take place" (1997, p.160). The emphasis on retrieval and storage is indicative of the narrow conception: the computational processes under investigation are localized within the individual. It is only once items are learned and internalized that computational processes can operate over them. The Bayesian matching procedure applies to items stored internally within an individual's episodic memory.

Consider, next, how Dienes' conceives of his model, he writes: "[L]awful behaviour may be produced by a connectionist network in which rules or hypotheses are not explicitly represented" (1992, p.40). A little later he writes: "the subject of the models obeys the rules, but does not represent them symbolically"(1992, p.70). Again, the message is plain. The artificial neural network represents a cognitive system that employs internal representations and rules that solve artificial grammar tasks, and the human data helps to reveal these internal computational processes and structures. The connectionist network is meant to represent the internal computational system within a subject that is used to carry out the cognitive task.

Finally, consider Osherson et al.'s model. In studying inductive reasoning, Osherson et al. adopt the following position: “The similarity-coverage model assumes that the existence of a pre-established hierarchy of categories that classify the instances figuring in an argument. The success of the model in predicting the qualitative phenomena…testifies to the approximate soundness of the model’s assumptions” (1993, p.200). What emerges, again, is a particular interpretation of what has been revealed about the underlying computational system. Reasoning about inference chains is an internal computational process that requires the deployment of particular categorical hierarchies. The boundary of the cognitive system is once again fixed at the formal system detecting relationships between argument stimulus input and subordinate categories.

Each of the three examples conforms, in varying degrees, to the narrow conception. The individual or some sub-component is the complete and natural unit of computational theorizing. The individual, in each case, is conceived of, and studied as if it were, the largest organized set of components capable of supporting computational investigation.

But notice that in addition to helping researchers to better understand models, the narrow conception also helps to structure the way in which researchers go about identifying and constructing investigations. The narrow conception also offers a means for thinking about where and what to look for when during investigation. It proposes methodological guidelines for studying computational cognitive systems.

Recall, for instance, that each of the three models addressed particular problems, proposed different solutions, and provided different explanations. Shiffrin and Stevyers’ model, for instance, conceived of recognition as a problem of item matching. This meant that the computational processes involved searching through memory traces using a global matching procedure. Dienes’ model, on the other hand, conceived of implicit learning as a form of pattern recognition. This led to looking for the internal exemplar representations and rules that allowed the network to identify and classify new letter strings. Finally, in Osherson et al.’s study, inference was taken to involve detecting structural category relations. This meant that it attempted to build a model around understanding how such categorical relationships could be structured.

One way to understand why each study offers the types of model it does and measures the fit of its model(s) against the types of experimental data that it does is as a result of the constraining influence of the narrow conception. In directing attention to the individual and its sub-components, the narrow conception sets up certain implicit conceptual boundaries. It limits which computational explanations are seen as viable, which properties and processes are taken to be necessary for investigation, and which solutions are considered plausible. The explanatory space of options surrounding computational theorizing is delimited. The narrow conception curbs the conceptual and methodological understanding of computation available for use within investigations.

The Wide Conception

The discussion up until this point has been largely descriptive. The goal has been to articulate what the narrow conception amounts to and provide a sense of the way in which it imposes interpretative and methodological constraints on research. In this final section, the aim is to provide a critical analysis of the view. Two problems are raised.

The first problem follows on the heels of the constraining influence of the narrow conception. The issue is that if the narrow conception limits the theoretical and explanatory horizons of computational investigations, then it also limits the kinds of research that can conducted. This is an undesirable state of affairs insofar as a healthy domain of investigation should have the broadest range of alternatives available when conducting research. If researchers are limited in the potential avenues they might explore, then the range of theories, explanations, and models they end up offering may turn out to be impoverished. In an ideal world, there will be as few constraining or biasing assumptions as possible during investigation. Insofar as the narrow conception operates as a constraining assumption on computational psychology, it forms a barrier to conducting successful research.

The history of behaviorism offers an instructive example. In both its logical and philosophical forms, behaviorism eschewed recourse to ‘mental’ vocabulary. It held that only ‘observable behaviour’ was the proper subject of psychological investigation. One result of its constraining influence was North American psychology made little reference to mental structures and processes. It took almost 30 years to reclaim the conceptual territory lost to behaviorism (Gardner, 1985). The claim here is not quite so negative, but the moral is the same. The narrow conception has potentially closed off interesting avenues of computational research because of its constraining influence.

One might respond by arguing that the above concern is only a really problem if the narrow conception turns out to be false. But that given the wealth of empirical support the view enjoys, there is really no reason to think that the narrow conception is in fact not the right view to hold. The problem with this response is that gets the order of explanation backwards. It is not that the narrow conception is true *because* computational research conforms to its strictures. Rather, it is because the narrow conception imposes certain restrictions on research that computational investigations conform to its strictures. The narrow conception problematically limits the range of alternatives considered before, during and after investigation.

The second concern is that the narrow conception, on occasion, provides explanatory weaker accounts of psychological phenomena in virtue of its over emphasis on individual-bound systems. Because the narrow conception emphasizes the individual as the limit of computational explanations, investigations based on its strictures can fail to identify the important computational role played by environmental elements.

Consider an example from the history of cognitive science. Problem solving was traditionally thought to involve a search through problem space (Newell, Shaw, & Simon, 1960). One way this approach was computationally instantiated was to simulate agents searching mentally through a virtual problem space during various tasks (Newell & Simon, 1976). One issue with these early approaches is that cognizers often interactively explore problems by physical manipulating external structures (Kirsh, 2009). These types of actions are more than just pragmatic, as they crucially help cognizers to simplify and transform complex problems. Computational models that focused narrowly on internal searches missed out on the simplifying computational role of epistemic actions (see Wilson, 2004; Clark, 2008).

Insofar as computational explanations fail to pay sufficient attention to elements of the environment that offload and distribute cognitive activities, they stand to provide weaker accounts of psychological phenomena. Computational explanations that are overly reliant on the narrow conception, such as in the above case, can supply explanatorily weaker accounts (Wilson, 2014). This is not to say that every computational explanation that subscribes to the narrow conception is explanatorily weaker. Rather, it is to point out that because there are blind spots imposed by the narrow conception, some computational explanations may, on occasion, be weaker than potential alternatives.

The previous two concerns should not be taken to undermine the narrow conception in its entirety. Instead, the concerns are better understood as forming a negative case against the sufficiency of the narrow conception as a global thesis. Given this, it will be worth exploring a possible alternative approach to understanding computation.

Wide computationalism is the idea that at least some of the elements of computational cognitive systems can reside outside the individual (Wilson, 1994, 1995; Hutchins, 1995; Kersten, 2016; Kersten & Wilson, 2016). Wide computational systems are those systems that recruit computational units from the larger embedding environment. Similar ideas have also been offered about cognition under the label of 'situated, embedded and extended' cognition (see Wilson, 2004; Clark 2008).

The viability of wide computationalism follows from the location neutrality of computational individuation. Wilson, for example, writes: "There is nothing in the method of computational individuation itself…which implies that the class of physical features mapped by a realization function cannot include members that are part of the environment of the individual" (1994, p.355). Because formal systems are medium neutral, it is at least possible that some of the computational elements include parts outside the individual. Wide computationalism stands in contrast to the narrow conception insofar as it pushes computational analysis outside the individual. Wide computationalism also gains support from a number of empirical studies in human and animal psychology (see Kersten, 2016).

Wide computationalism is a locational thesis about the realization or supervenience base of computational cognitive systems. It is a view about the scope of physical systems, processes, and components that are capable of supporting computational analysis. What this means is that although wide computationalism is compatible with either an individualist (Segal, 1991) or an externalist (Shagrir, 2001) interpretation, it is, strictly speaking, non-committal on issues of representational or semantic individuation.

For present purposes, the truth of wide computationalism is less important than the alternative it presents. This is because wide computationalism provides one potential alternative for understanding computation within computational psychology. In articulating a conception of computation that moves beyond the individual, wide computationalism stands to supply an importantly distinct approach to understanding computational investigations. By exploiting the location neutrality of computational individuation, wide computationalism re-conceptualizes the study of computational cognitive systems as at least partially requiring analysis of the embedding environment.

Investigations based on this wide approach stand to pay closer attention to the role of the environment, given their explicit focus on computational systems spreading out across the brain, body and world. Examples of the wide conception in action, for example, include agent-based models or swarm behaviour models (see Dawson, 2010). One way to view wide computationalism, then, is as an alternative conception of the underlying concept of computation that may be used within computational psychology.

Another way to make the point is to say that whereas the narrow conception might be construed as a restrictive monistic and a priori assumption about how cognitive states and processes are studied, wide computationalism provides an alternative pluralistic, empirical approach to investigation. Instead of viewing the narrow conception as exhausting the logical space of investigation, wide computationalism might be seen as a further, important additional explanatory strategy that can be used when thinking about computational investigations. Some phenomena may be more amendable to wide investigation, while others may conform more closely to the narrow conception. It may be that in some cases a narrow approach is preferable, while in others a wide approach is more suitable. In opening up the logical space, computational psychology is better positioned to precede both methodologically and theoretically.

This is only the briefest of sketches, but it should begin to provide a sense of how computational psychology may move beyond the narrow conception. However, the wide approach is not offered as a replacement to the narrow conception, but rather as a supplement. Wide computationalism is simply an extension of the logic inherent within computational psychology. The point is that it can step in when computational investigations run up against the limits of the narrow conception. On the proposed view, research that conforms to the narrow conception, such as the three examples surveyed, still makes a valuable and important contribution to cognitive science and psychology.

The general point to note in concluding is that in demonstrating the commitment of three paradigmatic examples of computational research to the narrow conception and outlining two problems the view faces, the case for the existence and problematic influence of the view has been at least partially motivated. The narrow conception has, on occasion, problematically structured at least some of the thinking within computational psychology, and that in doing so it has laid down some of the conceptual track on which the computational research train has run. Given this, further examination of previously underexplored approaches, such as wide computationalism, may help enrich the range of theories, models, and explanations offered within computational psychology.

References

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J., & Lebiere, C. (2003). The newell test for a theory of cognition. *Behavioral and Brain Sciences*, 26, 587–640

Anderson, J., & Liebere, C. (1998). *The atomic components of thought.* Mahwah, NJ: Lawrence Erlbaum Associates.

Dawson, M.R.W., Dupuis, B., & Wilson, M. (2010). *From Bricks to Brains: The Embodied Cognitive Science of LEGO Robots*. Edmonton: Athabasca University Press.

Boden, M. A. (1981). Minds and Mechanisms: Philosophical Psychology and Computational Models. Brighton, Sussex: The Harvester Press.

Clark, A. (2008). Supersizing the Mind: Embodiment, Action, and Cognitive Extension. Oxford: OUP.

Cleeremans, & Dienes, Z. (2008). Computational Models of Implicit Learning. In R. Sun (ed.), The Cambridge Handbook of Computational Psychology (pp.396-421). Cambridge, MA: Cambridge University Press.

Dienes, Z. (1992). Connectionist and memory- array models of artificial grammar learning. *Cognitive Science, 16*, 41–79.

Egan, M. F. (2000). Computation and Content. *The Philosophical Review*, 104(2), 181-203.

Fodor, J. (1983). *Modularity of Mind*. Cambridge, MA: Bradford/MIT Press.

Gardner, H. (1985). *The Mind's New Science*. New York: Basic Books.

Glanzer, M., Adams, J. K., Iverson, G. J., & Kim, K. (1993). The regularities of recognition memory. *Psychological Review*, *100*, 546– 567.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge: MIT Press.

Kersten, L. (2016). A Mechanistic Account of Wide Computationalism. *Review of Psychology and Philosophy*, (online), 1-17. DOI: 10.1007/s13164-016-0322-3.

Kersten, L., & Wilson, R.A (2016). The Sound of Music: Externalist Style. *American Philosophical Quarterly*, *53*(2), 139-154.

Kirsh, D. (2009). Problem Solving and Situated Cognition. In P. Robbins and M. Aydede, *The Cambridge Handbook of Situated Cognition* (pp. 264-306). Cambridge, Mass.: Cambridge University Press.

Milkowski, M. (2015). Computational mechanism and models of cognition. *Philosophia Scientiae*, 18(3), 1–14.

Newell, A. (1990). *Theories of Unified Cognition*. Cambridge, MA: Harvard University Press.

Newell, A., & Simon, H. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.

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

Osherson, D. N., Smith, E.E., Wilkie, O., Lopez, A., Shafir, E. (1990). Category-based induction. *Psychological Review*, 97, 185-200.

Piccinini, G. (2015). Physical Computation: A Mechanistic Account. Published to Oxford Scholarship Online. doi: 10.1093/acprof:oso/9780199658855.001.0001.

Pylyshyn, Z. (1984). *Computation and Cognition: Toward a Foundation for Cognitive Science*. Cambridge, MA: MIT Press.

Rowlands, M. (1999). The Body in Mind: Understanding Cognitive Processes. Cambridge: CUP.

Segal. G. (1991). Defence of a Reasonable Individualism. *Mind*, 10, 485-494.

Shagrir, O. (2001). Content, Computation, and Externalism. *Mind*, 110 (438), 369-400.

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM – retrieving effectively from memory. *Psychonomic Bulletin and Review*, 4, 145–166.

Simon, A. (1962). The Architecture of Complexity. *Proceedings of the American Philosophical Society* 106, 467-482.

Sternberg, S. (1969). Memory Scanning: Mental Processes Revealed by Reaction Time Experiments. *American Scientist*, 57, 421-457..

Sun, R. (2008). Theoretical status of computational cognitive modeling. *Cognitive Systems Research*, 1-17.

Sun, R., & Ling, C. (1998). Computational cognitive modeling, the source of power and other related issues. *AI Magazine*, *19*(2), 113–120.

Wilson, R. (1994). Wide computationalism. *Mind* 103(4): 351–372.

Wilson, R. (1995). *Cartesian psychology and physical minds: Individualism and the sciences of the minds*. Cambridge: Cambridge University Press.

Wilson, R. (2004). *Boundaries of the mind: The individual in the fragile sciences*. Cambridge: Cambridge University Press.

Wilson, R. (2014). Ten questions concerning extended cognition. *Philosophical Psychology* 27(1): 19–33.