

## Foundations of Computational Neuroscience

Gualtiero Piccinini

Oron Shagrir

Highlights:

- Neural systems process information by performing computations
- Neural computation and information processing operate at many levels of organization
- Computational neuroscience is an essential part of the science of cognition

**Abstract:** Most computational neuroscientists assume that nervous systems *compute* and *process information*. We discuss foundational issues such as what we mean by “computation” and “information processing” in nervous systems; whether computation and information processing are matters of objective fact or of conventional, observer-dependent description; and how computational descriptions and explanations are related to other levels of analysis and organization.

### 1. Introduction

Computational neuroscience has two faces. On one hand, it builds computational models of neural phenomena, analogously to the way computational chemistry, climate science, computational economics, etc. build computational models of their respective phenomena. On the other hand, computational neuroscience studies the way nervous systems *compute* and *process information*. Thus, unlike computational scientists in most other disciplines, computational neuroscientists often assume that *nervous systems* (in addition to the scientists who study them) perform computations and process information.

Consider for example the neural integrator that converts eye-velocity inputs to eye-position outputs, and thus enables the oculomotor system to move the eyes to the right position [1]. A variety of computational models have been offered for this network [2]–[6]. In addition, it is assumed that the integrator itself *processes information* about eye velocities and eye positions and produces eye-position codes by *computing* mathematical integration over these eye-velocity encoded inputs.

Is this assumption correct? That depends not only on what nervous systems do but also on what we mean by “computation” and “information processing”. This leads us into the foundations of computational neuroscience.

As to computation, there is a precise and powerful mathematical theory that defines which functions of a denumerable domain, such as the natural numbers or strings of letters from a finite alphabet, can be computed by following an algorithm. The same theory shows how to build machines that can compute

any function that is computable by algorithm—that is, universal computers [7]. Our ordinary digital computers are universal in this sense until they run out of memory.

But the mathematical theory of computation does not tell us whether and how nervous systems perform computations, and in what sense. This is because the mathematical theory of computation was never intended to be and indeed is not a theory of physical computation, namely, of physical computing systems such as brains. Thus there might be hypothetical physical systems that compute functions that are not Turing machine computable [8], [9]. Furthermore, there are many physical systems whose performance is described by computable functions even though we do not say that the systems compute the functions. A rock that is sitting still, for example, does not compute the identity function that describes some of its behavior (or lack thereof).

As to information, there is also a precise and powerful mathematical theory that defines information as the reduction of uncertainty about the state of a system. The same theory can be used to quantify the amount of information that can be transmitted over a communication channel [10]. Again, the mathematical theory of information does not tell us whether and how the brain processes information, and in what sense. So establishing the foundations of computational neuroscience requires more work.

Foundational discussion is important because it articulates the explanatory scope of computational descriptions, the relations between computational level and other levels of description (see Section 4) and the metaphysical commitments carried by the terms ‘information’ and ‘computation’. Take the oculomotor integrator. We say that it encodes information about eye velocities and positions and that it computes integration. Do we take this statement as a commitment to real, objective facts in the brain, or is it just a useful way to describe the brain used by scientist for heuristic or illustrative purposes? Churchland, Koch and Sejnowski [11], for example, state that "whether something is a computer has an interest-relative component, in the sense that it depends on whether someone has an interest in the device's abstract properties and in interpreting its states as representing states of something else" (p. 48). Others have replied that, on the contrary, whether something computes and processes information is an objective fact [12].

A related question concerns whether every physical object is a computer. Putnam [13] argues that every physical system satisfying minimal conditions implements every finite state automaton. Assuming that to compute is to satisfy Putnam’s minimal conditions, this implies that every physical object, including rocks and chairs, computes practically everything! (See also [14].) Many have replied that Putnam assumes a much too liberal notion of implementation (e.g., [15], [16]). Chalmers [17], for example, concedes that everything computes something, but insists that only few objects implement the kind of automata that suffice for minds (see [18] for further replies and discussion). Answering these questions depends on how we apply the notions of information and computation to physical systems.

## 2. What is information?

Let’s begin with information. There is no doubt that nervous systems contain internal variables that correlate reliably with other variables, both internal and external to it. For instance, neuronal spike

trains correlate reliably with other neuronal spike trains from other neurons and with aspects of the environment such as light, sound waves, pressure, and temperature.

This is enough to establish that nervous systems carry information in two senses [19]. First, they carry information in Shannon's sense—some of their variables reduce uncertainty about other variables. E.g., certain spike trains in the oculomotor system correlate reliably with eye movements. Information in Shannon's sense has to do with the uncertainty that characterizes a process as a whole, including all of the possible alternative messages at once. The Shannon information generated by the selection of a particular message is a function of how many alternative messages may be selected instead and the probability with which any possible message is selected.

By contrast, semantic information has to do with what a particular signal stands for or means. To capture the semantics of a signal, it is not enough to know which other signals might have been selected instead and with what probabilities. We also need to know what a particular signal stands for. Different equiprobable messages carry the same amount of Shannon information, but they may well mean completely different things. We call 'semantic information' the information a signal carries by reducing uncertainty about a specific state of affairs. Nervous systems carry semantic information in the sense that specific states of some of their variables make it likely that other variables (which they reliably correlate with) are in certain specific states. E.g., a certain spike train in the oculomotor integrator makes it likely that a specific eye movement is about to occur.

Our opinion is that at least some neural variables carry information in a third sense too—the sense in which neural variables *represent* the environment as being a certain way. Representation is something more than mere semantic information (which in turn is something more than Shannon information). This is because representation can be either correct or incorrect (in which case it is a *misrepresentation*), whereas mere semantic information, by itself, is neither correct nor incorrect (either a signal raises the probability of a state of affairs or it doesn't; there is nothing right or wrong either way). In this third sense of information, neural events are not merely correlated with a state of the world but *represent* such a state of the world, which means that they may be either correct or incorrect about how the world is. For instance, let's assume that there are neural events in every speaker's Wernicke's area corresponding to each utterance. Some neural events correspond to true utterances such as "The Moon is a satellite of the Earth." Those neural events truly represent a state of the world, e.g., that the Moon is a satellite of the Earth. Other neural events correspond to false sentences such as "the Martians have invaded the Earth." Those neural events *misrepresent* the world as different than the way it is.

There are those who think that *neural* representation, as neuroscientists understand it, is insufficient for genuine *mental* representation—that is, the kind of representation that we usually attribute to each other's minds (beliefs, desires, mental images, etc.) [20]-[22]. Others think that neuroscience already assumes a notion of representation even stronger than the one we just mentioned, to be discussed below [23], [24].

### 3. What is physical computation?

Let us turn to computation. Some philosophers have tried to explain what it takes for a physical system to perform computations by using notions found in logic and computability or automata theory. They describe computation as program execution [25], syntactic operations [26], [27], automatic formal systems [28], or implementation of automata [17]. These notions might apply to digital computers. But, as many have noted, the brain is very different from the familiar digital computers [29]-[34]. In nervous systems, the functional relevance of neural signals depends on non-digital aspects of the signals such as firing rates and spike timing. Therefore, there is a strong case to be made that typical neural signals are not strings of digits, and neural computation is not, in the general case, digital computation [35].

Two more recent views about computation in the brain reflect the authors' (somewhat opposing) opinions. According to the modeling view of computation, physical computation is a special form of representation – it is a dynamic *model* in the sense that it represents a target domain in a way that preserves its high-order structures [36]-[38] (The modeling view is stronger than the semantic view that merely identifies computation with information-processing [39].) The oculomotor integrator, for example, preserves the integration relation between velocity and positions of the (represented) eyes. The distance between two successive (encoded) eye positions is just the integration over the (encoded) eye velocity with respect to time. The claim is that computational neuroscience often invokes this isomorphism-based, stronger notion of representation (i.e., computation).

To elaborate a bit, the oculomotor integrator is a representational system in the sense that its cells encode information about eye velocities and eye positions. But, in addition, the oculomotor integrator is a computing system in the sense that it preserves a high-order, mathematical, integration relation between eye movement and eye positions. By computing integration, this oculomotor system mirrors or preserves the integration, movement-position, relation. According to this view, computation has nothing to do with the mechanism by which the system computes integration. This does not mean that computational neuroscientists should not characterize computing mechanisms; of course they should and do so. The claim is that the neural integrator computes regardless of the mechanism that carries out the computation. The mechanism could be a digital process, but it could also be a dynamical system operating on continuous variables, such as the line-attractor network proposed for the oculomotor integrator [2]-[4]. Rather, the mechanism is a computing one because it models high-order structures in the target, represented, domain in a way that preserves such structures.

According to the other, mechanistic view, computation and information processing/representation are distinct notions [19]. Computation is a specific kind of mechanistic process; it has to do with the processing of variables to obtain certain relationships between inputs, internal states, and outputs independently of how the variables are physically implemented, and this is so regardless of whether the variables carry any information about the environment [40, 41].

In terms of our example, the mechanistic view agrees that the oculomotor integrator is a computing system, but not because it processes information or builds models (though it may well do that too). Rather, the oculomotor integrator performs computations because it manipulates certain internal variables so as to obtain specific functional relationships between its inputs and outputs, where the inputs and outputs are signals that enter and exit the oculomotor system and are characterized in terms

of differences between different portions of the signals (e.g., spike frequencies) rather than specific physical properties of the signals such as Calcium or Potassium ions flowing in and out of axons [19].

In spite of some conceptual differences between different authors, there is consensus that nervous systems process information by performing computations and that computation must be characterized by abstracting away from certain aspects of a physical system (from the implementing media for the mechanistic view; from the mechanisms themselves for the modeling view). The next question is, which kind of computations? The traditional answer is that nervous systems perform *digital* computations, like those performed by our artificial digital computers [42]. Some authors have gone beyond this to claim that in order for nervous systems—even rather simple systems, such as those of ants and bees—to carry out their cognitive functions, they must be functionally organized like digital computers, with a general purpose processor and an addressable read-write memory that can store strings of symbols [43]. But the theory that neural computations are digital is held by a small and shrinking minority. Few if any neuroscientists subscribe to it. The reasons are several, though they are not as straightforward as they may seem. We'll look at a couple of bad reasons and then get to some better ones.

For starters, the problem is not that nervous systems are massively parallel while digital computers are serial. That argument, which is often given, is a fallacy based on confusion between levels of organization [44]. Digital computers are serial at the level of their processor, but the processor itself is composed of hundreds of millions of logic gates that act in parallel. In addition, digital computers may contain multiple processors working in parallel. True, nervous systems appear to perform parallel computations in ways that are different from those of digital computers and that we do not yet fully understand, but parallelism by itself is not the problem.

Another bad reason is the alleged inseparability of processor and memory in nervous systems. While many neural networks perform both memory and processing functions, it may well be that at least in some cases, nervous systems perform processing and memory functions in separate subsystems. So the distinction between processor and memory is not the problem either.

But the theory that nervous systems are digital computers does face a serious problem, that is, the nervous system's primary computational vehicles—spike trains—are irreducibly graded in their functional properties. In other words, the functional relevance of neural signals depends on non-digital aspects of the signals such as firing rates and spike timing. Therefore, typical neural signals are not strings of digits, and neural computation is not, in the general case, digital computation.

This is not say that neural computations are analog. Strictly speaking, analog computation employs continuous signals, whereas neural signals are made out of discontinuous functional units—neuronal spikes or action potentials. Thus, neural computation appears to be neither digital nor analog; it appears to be a distinct kind of computation (Figure 1) [35].

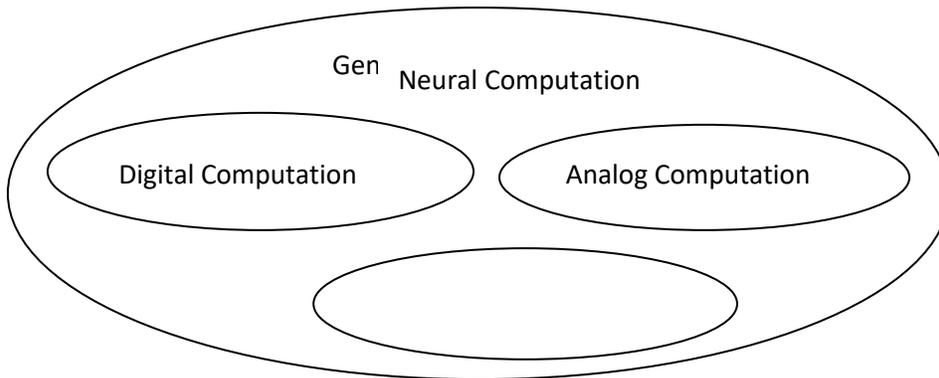


Figure 1. Some types of generic computation. Neural computation may sometimes be either digital or analog in character; but, in the general case, neural computation appears to be a distinct type of computation.

#### 4. Levels of organization and levels of analysis

Nervous systems as well as artificial computational systems have many levels of mechanistic organization [45], [46]. They contain large systems like the brain and the cerebellum, which decompose into subsystems like the cortex and the brainstem, which decompose into areas and nuclei, which in turn decompose into maps, columns, networks, circuits, neurons, and sub-neuronal structures. Computational neuroscience studies neural systems at all of these mechanistic levels, and then it attempts to discover how the properties exhibited by the components of a system at one level, when they are suitably organized into a larger system, give rise to the properties exhibited by that larger system. If this process of linking explanations at different mechanistic levels is carried out, the hoped result is an integrated, multi-level explanation of neural activity.

But computational neuroscience also involves levels of analysis. First, there is the level of what a neural subsystem does and why. Does it see or does it hear? Does it control the arm or the head? And what function does it compute in order to perform this function? Answering these *what* and *why* questions leads to what Marr called a “computational theory” of the system [47]. The theory specifies the function computed and why it is computed, without saying what representations and procedures are used in computing it. Specifying the representations and procedures is the job of the “algorithmic theory.”

Finally, an “implementation theory” specifies the mechanisms by which the representations and algorithms are implemented [47].

There is a debate about the role of the computational level and how it relates to the algorithmic and the implementation levels. Some authors argue that the computational analysis specifies the computed mathematical function [48], [49]. A variant of this view has been advertised by some proponents of Bayesian optimality analysis, who argue that their “focus is on computational-level theories, characterizing the functional capacities of human inference rather than specific psychological processes that implement those functions” [50, p. 206]. They further say that what is distinctive about Bayesian approach is “a top-down, or ‘function-first’ research strategy, as recommended by Marr (1982): from computational, to algorithmic, to implementational levels” [51, p. 196].

Others argue that the distinctive feature of Marr’s computational level is in anchoring the computed function in the individual’s environment [52], [53]. Marr implicitly assumes that the brain models the environment, in the sense—mentioned above—of preserving certain mathematical relations between environmental variables, e.g., variables in the visual field. Thus, for example, the zero-crossings of the second derivatives computed by early visual processes mirror the sharp changes in the reflection function that might occur along object boundaries. The role of the computational level is to specify the mathematical function that is being computed (zero-crossings of the second derivatives), and to demonstrate that this function preserves certain mathematical relations between environmental variables in the visual field (object boundaries). Demonstrating this, we can explain why computing differentiation (i.e., obtaining derivatives), and not, say, factorization or exponentiation, is appropriate for edge-detection. By the same token, the computational level explains why computing integration (and not say subtraction or differentiation) is appropriate for eye movements; the reason being that integration mirrors the eye’s velocity-position (integration) relation.

Other authors argue that computational explanation is fully mechanistic—computational and algorithmic theories are sketches of mechanisms. They are just partial descriptions of neurocomputational mechanisms at one or more mechanistic levels [54], [55]. In terms of our example, the oculomotor integrator is a computing mechanism. Even if we limit ourselves to a computational theory, which describes the function it computes (integration of eye-velocity inputs to eye-position outputs), we still need to make reference to the kinds of inputs and outputs being manipulated (say, firing rates), and those are concrete aspects of the mechanism that require at least a partial understanding of the components. If we go beyond a computational theory and search for a *correct* algorithmic theory (as opposed to a hypothesis about what algorithm may be in place), we need to know the different components of the mechanism, how they are connected, and what operations they perform. By the time we have enough details about the operations of the components of the mechanism to establish the correct algorithmic theory, we are well on our way to understanding how the algorithm is implemented. In summary, the three Marrian levels—computational, algorithmic, and implementational—are interdependent aspects of a description of a mechanism

Finally, the holy grail of neuroscience is explaining the mind or at least its cognitive aspects. Traditionally, many psychologists and philosophers maintained that neuroscience is only concerned with

implementing mechanisms, whereas the mind/cognition is the proper domain of psychology [56]. But psychology itself is increasingly turning into cognitive and computational neuroscience. And the kind of explanations provided by psychologists, when they are examined closely, turn out to be partial aspects of the kind of multi-level mechanistic explanations pursued by neuroscientists. How to integrate computational explanations in psychology and neuroscience is still an on-going heated debate that is constantly affected by new advances in neuroscience [57], [58].

#### References:

- 1) Robinson, D. A. 1989. Integrating with neurons. *Annual Review of Neuroscience*, 12: 33–45.
- 2) Cannon, S. C. & Robinson, D. A. 1985. An improved neural-network model for the neural integrator of the oculomotor system: More realistic neuron behavior. *Biological Cybernetics*, 53: 93–108.
- 3) Seung, S.H. 1996. How the brain keeps the eyes still. *Proceedings of the National Academy of Science USA*, 93: 13339–44.
- 4) Eliasmith, C. & Anderson C. H. 2003. *Neural Engineering: Computation, Representation and Dynamics in Neurobiological Systems*. Cambridge, MA: MIT Press.
- 5) Loewenstein, Y. & Sompolinsky, H. 2003. Temporal integration by calcium dynamics in a model neuron. *Nature Neuroscience*, 6: 961–967.
- 6) Goldman, M.S., Compte, A. & Wang, X.J. 2009. Neural integrator models, In L.R. Squire (ed.), *Encyclopedia of Neuroscience* (vol. 6). Oxford: Academic Press, pp. 165-178.
- 7) Turing, A.M. 1936. On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, 42: 230-265.
- 8) Copeland, B.J. & Shagrir, O. 2007. Physical computation: How general are Gandy's principles for mechanisms? *Minds and Machines*, 17: 217-231.
- 9) Piccinini, G. 2011. The physical Church-Turing thesis: Modest or bold? *British Journal for the Philosophy of Science*, 62: 733-769.
- 10) Shannon, C.E. 1948. *A Mathematical Theory of Communication*. Bell System Technical Journal, 27: 379–423, 623–656.
- 11) Churchland, P.S., Koch, C. & Sejnowski T.J. 1990. What is computational neuroscience? In E.L. Schwartz (ed.), *Computational Neuroscience*. Cambridge, MA: MIT Press, pp. 46-55.
- 12) Piccinini, G. 2007. Computational modeling vs. computational explanation: Is everything a Turing machine, and does it matter to the philosophy of mind? *Australasian Journal of Philosophy*, 85: 93 – 115.
- 13) Putnam, H. 1988. *Representations and Reality*. Cambridge, MA: MIT Press.
- 14) Searle, J.R. 1992. *The Rediscovery of the Mind*. Cambridge, MA: MIT Press.
- 15) Chalmers, D.J. 1996. Does a rock implement every finite-state automaton? *Synthese*, 108: 309-333.
- 16) Godfrey-Smith, P. 2009. Triviality arguments against functionalism. *Philosophical Studies*, 145: 273-295.
- 17) Chalmers, D.J. 2011. A computational foundation for the study of cognition. *Journal of Cognitive Science*, 12: 323-357.

- 18) Piccinini, G. (ed.) 2011-2012. Special issues on [17] including 13 papers by different authors and Chalmers' reply. *Journal of Cognitive Science* 12(4) – 13(1-3).
- 19) Piccinini, G. & Scarantino, A. 2011. Information processing, computation, and cognition, *Journal of Biological Physics*, 37:1-38.
- 20) Bennett, M. R. & Hacker, P. M. S. 2006. Language and cortical function: Conceptual developments. *Progress in Neurobiology*, 80: 20–52.
- 21) Ramsey, W. 2007. *Representation Reconsidered*. Cambridge: Cambridge University Press.
- 22) \*Shea, N. 2013. Naturalising representational content. *Philosophy Compass*, 8: 496-509. Argues for pluralism about representational content and that data from cognitive neuroscience should play a greater role in theories of content.
- 23) Sprevak, M. 2011. Review of W. M. Ramsey, *Representation Reconsidered*. *The British Journal for the Philosophy of Science*, 62: 669-675.
- 24) Shagrir, O. 2012. Structural representations and the brain. *British Journal for the Philosophy of Science*, 63: 519-545.
- 25) Cummins, R. 1989. *Meaning and Mental Representation*. Cambridge, MA: MIT Press.
- 26) Pylyshyn, Z.W. 1984. *Computation and Cognition*. Cambridge, MA: MIT Press.
- 27) Fodor, J.A. 2008. *LOT 2: The Language of Thought Revisited*. Oxford: Oxford University Press.
- 28) Haugeland, J. 1981. Semantic Engines: An Introduction to Mind Design. In J. Haugeland (ed.), *Mind Design*, Cambridge, MA: MIT Press, pp. 1-34.
- 29) \*\*Von Neumann, J. (2012/1958). *The Computer and the Brain*. London: Yale University Press (The 3<sup>rd</sup> edition includes a foreword exposition by Ray Kurzweil).
- 30) Churchland, P.S. & Sejnowski, T. 1992. *The Computational Brain*. Cambridge, MA: MIT Press.
- 31) Churchland, P.M. 1989. *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*. Cambridge, MA: MIT Press.
- 32) Bechtel, W. & Abrahamsen, A. 2002. *Connectionism and the Mind: Parallel Processing, Dynamics, and Evolution in Networks* (2<sup>nd</sup> edition). Oxford: Basil Blackwell.
- 33) Churchland, P.M. 2007. *Neurophilosophy at Work*. Cambridge: Cambridge University Press.
- 34) Edelman, S. 2008. *Computing the Mind: How the Mind Really Works*. Oxford: Oxford University Press.
- 35) \*Piccinini, G. & Bahar, S. 2013. Neural computation and the computational theory of cognition. *Cognitive Science*, 37:453-488. Argues that neural computation is neither digital nor analog but sui generis.
- 36) Grush, R. 2001. The Semantic challenge to computational neuroscience, in P. Machamer, R. Grush & P. McLaughlin (eds.), *Theory and Method in the Neurosciences*. Pittsburgh: University of Pittsburgh Press, pp. 155-172.
- 37) Shagrir, O. 2006. Why we view the brain as a computer. *Synthese*, 153: 393-416.
- 38) \*Shagrir, O. 2010. Brains as analog-model computers. *Studies in History and Philosophy of Science*, 41: 271–279. Argues that neural computation consists in constructing models of the environment.
- 39) Sprevak, M. 2010. Computation, individuation, and the received view on representation. *Studies in History and Philosophy of Science* 41: 260-270.
- 40) Piccinini, G. 2007. Computing mechanisms. *Philosophy of Science*, 74: 501-526.
- 41) Miłkowski, M. 2013. *Explaining the Computational Mind*. Cambridge, MA: MIT Press.
- 42) McCulloch, W.S. & Pitts, W. 1943. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5: 115-133.
- 43) Gallistel, C. R. & King, A. P. 2009. *Memory and the Computational Brain: Why Cognitive Science will Transform Neuroscience*. New York: Wiley/Blackwell.
- 44) Piccinini, G. 2008. Computers. *Pacific Philosophical Quarterly*, 89: 32-73.

- 45) Craver, C. F. 2007. *Explaining the Brain: Mechanisms and the mosaic unity of neuroscience*. New York: Oxford University Press.
- 46) \*Bechtel, W. 2008. *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience*. London: Routledge. A philosophical analysis of scientific, mechanistic, accounts of mental activities.
- 47) \*\*Marr, D. C. 2010/1982. *Vision: A computation Investigation into the Human Representational System and Processing of Visual Information*. San Francisco: Freeman. Articulates the distinction between the computational, algorithmic, and implementation levels of analysis. The 2010 edition (MIT) includes foreword by Shimon Ullman and afterword by Tomaso Poggio that discuss the relevance of Marr's approach today.
- 48) Egan, F. 2010. Computational models: A modest role for content. *Studies in History and Philosophy of Science*, 41: 253-259.
- 49) \*Egan, F. 2014. Function-theoretic cognitive explanation and neural mechanisms, in D. Kaplan, (ed.) *Integrating Mind and Brain Science: Mechanistic Perspectives and Beyond*. Oxford University Press (forthcoming). Argues that computational descriptions are genuinely explanatory even when resist a transparent mapping to known neural mechanisms.
- 50) Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. 2006. Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Science*, 10: 309-318.
- 51) Chater, N., Goodman, N., Griffiths, T. L., Kemp, C., Oaksford, M., & Tenenbaum, J. B. 2011. The imaginary fundamentalists: The unshocking truth about Bayesian cognitive science. *Behavioral and Brain Sciences*, 34:194-196.
- 52) Shagrir, O. 2010. Marr on computational-level theories. *Philosophy of Science*, 77: 477-500.
- 53) Bechtel, W. & Shagrir, O. 2013. The Non-Redundant Contributions of Marr's Three Levels of Analysis for Explaining Information Processing Mechanisms. *TopiCS* (Topics in Cognitive Science), forthcoming.
- 54) Kaplan, D.M. 2011. Explanation and description in computational neuroscience. *Synthese*, 183:339-373.
- 55) \*Piccinini, G. & Craver, C.F. 2011. Integrating psychology and neuroscience: functional analyses as mechanism sketches. *Synthese*, 183: 283-311. Argues that traditional psychological explanations are mechanism sketches that can be integrated with the multi-level mechanistic explanations provided by neuroscience.
- 56) Weiskopf, D.A. 2011. Models and mechanisms in psychological explanation. *Synthese*, 183: 313-338.
- 57) Peebles, D. & Cooper, R.P. (eds.) 2013: *The Role of Process-Level Theories in Contemporary Cognitive Science*. *TopiCS* (Topics in Cognitive Science), forthcoming.
- 58) Kaplan, D., (ed.) 2014. *Integrating Mind and Brain Science: Mechanistic Perspectives and Beyond*. Oxford University Press (forthcoming).