

WHEN WEAK MODULARITY IS ROBUST ENOUGH?

MARCIN MILKOWSKI

Institute of Philosophy and Sociology. Polish Academy of Sciences, Warsaw
marcin.milkowski@obf.edu.pl

Abstract

In this paper, I suggest that the notion of module explicitly defined by Peter Carruthers in *The Architecture of The Mind* (Carruthers 2006) is not really in use in the book. Instead, a more robust notion seems to be actually in play. The more robust notion, albeit implicitly assumed, seems to be far more useful for making claims about the modularity of mind. Otherwise, the claims would become trivial. This robust notion will be reconstructed and improved upon by putting it into a more general framework of mental architecture. I defend the view that modules are the outcome of structural rather than functional decomposition and that they should be conceived as near decomposable systems.

KEY WORDS: Modularity; Carruthers; Decomposition

Resumen

En este trabajo, sugiero que la noción de módulo explícitamente definida por Peter Carruthers en *La Arquitectura de la Mente* (Carruthers 2006) no se usa realmente en el libro. En su lugar parece adoptarse una noción más robusta. Esta noción más robusta, aunque asumida implícitamente, resulta mucho más útil para poder formular afirmaciones sobre la modularidad de la mente. De otro modo, las afirmaciones resultarían triviales. Esta noción robusta será reconstruida y mejorada por medio de su ubicación en un marco más general de la arquitectura mental. Defiendo la idea de que los módulos son el resultado de una descomposición estructural y no funcional, y que deben ser concebidos como sistemas casi descomponibles.

PALABRAS CLAVE: Modularidad; Carruthers; Descomposición

Weak modularity

In most discussions about modularity, it has become almost ritual to start with Jerry Fodor's notoriously strict notion of a mental module (Fodor 1983). Yet, showing in how many ways you can deny Fodor's claims is not really helpful in defining the proper notion of modularity. I suggest we should start with a much weaker notion of modularity such as suggested by Carruthers (2006), and then make it more robust by seeing how modularity is really used, both in philosophy and in science, especially in evolutionary psychology.

The notion that Carruthers recommends is a very weak one—it is almost a common-sense notion of module, understood roughly the way we understand modules in home appliances such as DVD players. A module is accordingly a functional subcomponent, and a subcomponent is a part dissociable from the rest of the system (Carruthers 2006, 18; a similar notion is defended by Barrett & Kurzban 2006). Functions of modules are being analyzed in two ways: in task analysis, and in performance analysis (relative speed of various operations indicates how the system is built). Modules are not distinct, monadic objects; they are rather interconnected and relatively isolated components of the same system. Therefore, a weak massive modularity claim will be that the mind is composed of a great many functionally distinct components.

But all this is on the verge of being trivial. Even harsh critics of modularity can accept such a weak notion (Prinz 2006). Unless you can show that the function of the modules meant here is not really Cummins-style function (Cummins 1975), the upshot is that any system with many causally active parts seems to be massively modular. Cummins-style function is only a causal role that contributes to capacities of a system, and all subcomponents of the system clearly play such roles (trivially, they contribute to the system-wide capacity of the system to include such subsystems). Thus, if there are many subcomponents in the brain, such as neurons, then the weak modularity thesis is true. This victory is, however, Pyrrhic; there is no substantial claim to vindicate, and just anything in the brain is a module, even a single neuron or a single atom.

In reality, much more is at stake; not everything is a subcomponent, only some capacities count as distinct functions, and, in general, a more robust notion of module should be used, if we want the modularity thesis to be minimally interesting. A closer look at further chapters and introductory remarks in Carruthers' book reveals that such a notion is in fact implicitly used.

Let me point to some examples. The requirement that a module be a relatively stable subcomponent makes some causal processes irrelevant in the regular architecture of mind. In other words, only recurring processes count (Carruthers 2006, 56-57). Brain lesions indicate there are different modules realizing certain similar functions (Carruthers 2006, 127); in other words, it is not only task and performance analysis that is being used for discovering modules. Moreover, mental modules are analyzed in the light of evolutionary history, including histories of other species (humans are supposed to share mindreading and forms of learning with great apes; cf. Carruthers 2006, 152).

These and other methodological remarks are interspersed in the book. It could therefore seem quite tedious a task to compile a list of criteria used for individuation of modules, to say nothing about creating a systematic one. Yet, the view that is being defended seems to be deeply seated within biological concepts of modularity, and as it has been shown several times, biological modules are analyzable in terms of Herbert Simon's architecture of complexity (Simon 1996, 183-215; Callebaut & Rasskin-Gutman 2005). Simon's argument for modularity is also endorsed by Carruthers (1996, 12-13), as are his views on hierarchical organization of modular systems. It is surprising that his reliance on Simon stops there, at least at the explicit level, while at the same time he seems to use the concept of modularity in a similar way. I will offer an account that is inspired by Simon's notion of a near decomposable system. A robust notion of module as defined below seems to be the one that is implicitly referred to in many discussions of modularity. The notion is a result of a rational reconstruction rather than of a tedious literal interpretation of various claims from different camps about architecture of mind, as they would be hardly reconcilable.

Near decomposable systems

In Simon's terminology, modules are parts of near decomposable systems, i.e., systems that contain relatively isolated subsystems.¹ Note that those subsystems are not required to be completely self-contained—nor are mental modules. Nobody supposes that any mental module implemented in a fragment of a brain could work without the rest of it. Subsystems are located within hierarchies, and can build complicated systems. This is intuitive when we look at “kludgy” biological systems. Carruthers seems to appreciate this point as well.

Simon's analysis is different from Fodorian and functional accounts of modularity (Bechtel, in press) because of the fact that subsystems are never meant to be completely encapsulated but rather relatively isolated. The permeable boundaries between modules depend on the relative frequency of interactions; inside the module, the interactions are far more frequent than with the rest of the system. The system is relatively isolated from the environment just in case it interacts with it less frequently than with its own parts (and without this isolation,

¹ Simon distinguishes also completely decomposable systems, but in biological world it seems that such systems do not exist—interconnections and dependencies are much more complex than in artificial hierarchies.

it is not a system at all). Which interactions are important here is determined by the structure of the module that implements its function. Though Simon does not mention other statistical features apart from relative frequency differences, other statistical measures could be deployed to track relevant causal chains.² The main difference with the functional account is that modules' identity is not determined merely by the functional analysis—which is nothing more than a theoretical ascription in Carruthers' book—but by their structure and position in the system. In other words, modules in this account are effects of structural rather than functional decomposition.

This does not mean that modules are not functional. They are functional but by the virtue of having proper structure and position in the organization of the system. Functions are not themselves modular, it is architectural subsystems that are relatively isolated and therefore modular. Positing modules as functional subsystems makes little sense if their identity depends wholly on their functionality, as we already have a notion of function; if there is nothing interesting to say about the structure that supports the function, then the notion of module is simply reducible to the notion of function and should be discarded from the theory as redundant and terminologically confusing. Many of the problems with defending the notion of functional modularity (as discussed by Barrett & Kurzban 2006) dissolve as theoretical artifacts as soon as the notion is replaced with an architectural one.

Structural modularity implies that components are not atomic black boxes—inspecting the interactions inside seems the best way to research systems. After all, this is what brain imagining technologies promise and try to achieve: they should give at least some insight into what interacts, in what way, how often, and what follows what. Modules are not static structures, they are interacting structured processes. In the case of mental modules, the frequency (if any) of neuronal connections indicates structural boundaries at a causal level. This approach, however hard to put into practice, seems far more *realist* and non-behaviorist. Realistic modeling is insofar crucial as mental modules are supposed to

² It is obvious that any physical system on Earth interacts with this celestial body all the time because of the gravitational force so this interaction would count as the most important according to a literal reading of Simon's definition. However, as it is the case for all systems, it can be ignored safely for most of them, *ceteris paribus*, as a statistically irrelevant constant. Moreover, some kinds of interrelations, as discovered by co-occurrence measures, could pop out only after applying more complex statistical measures. In the rest of the paper, I use the term "frequency" for the sake of verbal simplicity but I imply other kinds of statistically relevant interconnections.

be natural kinds—carved at the joints of mind, and supporting at least some specific regularity that helps them play an explanatory and predictive role. Modules that are explanatory and predictively useless have no place in an architecture of the mind—it makes no sense to posit them as candidates for basic structures of the mind.

What is missing both in Carruthers' and Simon's accounts is the double meaning of the system hierarchy. In the first case, specialized subsystems can be said to occur at the same level of system organization, e.g., different modules that take care of mindreading can be related hierarchically. Take a simple system such as a portable calculator. It has three main components: an input device, which is a keyboard, an output device, which is a display, and a processing unit. Consequently, subcomponents of these three components have hierarchically organized subparts. When trying to see why the keyboard does not work, most often it does not make sense to disassemble the display unit. It is just another hierarchy, and the hierarchic organization of the calculator is the very condition of its being serviceable.

In the second case, we speak about the hierarchy of organization levels. This is to say that lower level interactions subserve higher-level systems (Wimsatt 2000). Here, by organization levels, I mean roughly what is meant by ontological levels. William Wimsatt defines organization levels as local maxima of regularity and predictability in a multi-dimensional phase space of different modes of organization of matter. Wimsatt's complicated definition is often taken to mean roughly that it is the scale of objects that determines the ontological levels. However, this reading would make the definition invalid (Bechtel 2007): gravitational forces operate between objects of extremely different sizes. What is, however, meant by Wimsatt is rather that if you use a multi-dimensional space to describe objects, there will be some local maxima of regularity—that is, some parameters will be in a non-stochastic order. If you discover that you enter the local maximum, it means that you enter another organizational level. (Of course, this is not the end of the story, but it has less to do with modularity, so I'm leaving other interesting details aside.)

The multi-level organization of modular systems is an important point. For example, neurological evidence (brain lesions or brain imagining) for the existence of different processing centers, or mechanisms in the brain is clearly a case where lower-level properties are used to stipulate higher-level mental modules. This is important because of a methodological principle: if there is no reason to believe there is a lower-level implementation of higher-level modules, then the higher-level modules would be no modules at all. So this principle is used to confirm

that there are really some mental modules, and not just some *a priori* theoretical posits. Without the principle, any functional description of a mental capability could be taken as a functional description of a module, and that would allow us to posit more modules than we could ever count and want. For example, a higher-level capacity to learn to use a black ballpoint pen could be taken as different from the capacity to learn to use a white ballpoint pen (especially in dark places). It is clear that higher-level capacities could be actually results of interactions between lower-level functions. While it is harder to grab a black pen in a dark room (after all, you hardly can see it) than to grab a white pen, it does not mean that hand-writing is any different in both cases. So the requirement is not only that there is a lower-level implementation but that this implementation is not simply a result of incidental interaction between separate modules and functions. (This should not be read as a principle of greedy reductionism but rather as a principle that there can be no higher-level skyhooks posited as modules.)

There are however two problems left: even if we individuate modules based on the frequency of interactions, and look at lower-level processes, we must know what we mean by “functions of subsystems”. I do not think that we could simply say that what we need is Millikan proper functions (Millikan 1984). Proper functions or other functions defined in terms of evolutionary history, though sometimes explicitly referred to in vindications of functional decomposition (Barrett & Kurzban 2006), are not geared well toward new and novel functions which are individual, as new types of mechanisms are *ex definitione* dysfunctional on this account; this concerns also higher-level mechanisms acquired by individual learning (for other arguments against etiological notions of function, see Krohs 2007). A robust notion of function that does not boil down to functions of previous tokens, capacities or causal roles seems to be more suited to the task of functional analysis of modules. As I have already mentioned, Cummins-style function is dangerous for this purpose as it could make the whole modular account rather trivial.

Thus, since the structural modules are not the results of functional decomposition, the notion of function is not critical for defining them: they have their function in virtue of having a proper organization. The notion of function that is needed should be as realistic as to make functional ascription dependent upon the system structure and not on the theorist’s knowledge. For example, in Krohs 2004 account, a component of a system has a function if and only if it contributes to the system’s capacity to X, and having this function by the component should be a result of some selection process of possible system elements (prominently,

natural selection or a designer's decision). This kind of notion of function seems to be better suited for modular analyses (another useful notion for analyzing biological function is autonomy-based; see Colier 2002 or Bickhard 2000). I can remain relatively neutral about the specific notion of function used here insofar as it fulfills the basic requirements: (1) accounting for new functions, (2) distinguishing dysfunctions, (3) identifying multiple functions of the same structural subsystems and (4) delineating non-functional parts in systems. A more important problem is how to systematically check if there really is a module that plays a function.

Is that really a module?

Tooby and Cosmides (1997, 139) quote 12 re-engineering heuristic questions used to discover modules:

- 1 *Existence*. Does the module exist or is the function explainable by other (or more general) modules?
- 2 *Scope*. What is the content domain?
- 3 *Proper cognitive description*. What are procedures and representations in the module?
- 4 *Adaptive function*. Does it solve any adaptive problems?
- 5 *Universality*. Does it develop in all humans?
- 6 *Ontogenetic timing*. Is there a regular ontogenetic schedule of development?
- 7 *Activation*. When does the module run?
- 8 *Regulation and function*. What is regulated by the module, and what depends on it?
- 9 *Inter-relationships*. What role does it play in the computational network?
- 10 *Neural basis*. Is the module associated with a specific brain region?
- 11 *Role in real-world events*. Does it explain real-world (vs. laboratory) phenomena?
- 12 *Health implications*. Does malfunction of module play any role in clinical disorders?

This list is far from being systematic and exhaustive. But we can easily see a pattern here: there are heuristics related to the role of the module in the system (1, 8 and 9); heuristics related to discovery of stable universal structures (4, 5, 6, 11); heuristics related to specifying the

function of the module in computational terms (2, 3); heuristics related to lower levels of organization (10, 11), and to timing (6, 7).

Some of those heuristics are too strict. For example, there could be several similar modules in biological systems, and unless the function specification is very fine-grained, we could not discover them: take vision and the ventral/dorsal pathways (Milner & Goodale, 1996) as a clear example... Moreover, functions of some modules may overlap (and this could explain the brain flexibility, at least in part), and modules may serve more than one function. And there could be functions which are realized only by whole systems. Take blood circulation: this function is realized by the whole blood circulation system, not just the heart. Tooby and Cosmides make an unrealistic assumption of one-to-one correspondence between structures and functions (Carruthers 2006, 212 makes roughly the same assumption, but with more caveats). What should be expected (and what fuels much of the criticism in Prinz 2006) is that many mental mechanisms have multiple functions. This is the way biological systems are organized.

For the same reason, a module should not be ascribed a scope or content domain. It is the function of the module that has a scope, not the module itself. Mental modules are usually posited at the computational level, so their functions are understood as computational functions. It is obvious that any computation has an input and that physically realized computers or computational mechanisms react only to some kinds of inputs. In this sense, computational mental modules, as Barrett and Kurzban 2006 note, always have a restricted domain. Yet in this sense, domain specificity of modules is completely trivial, as even non-modular computational architectures would be domain specific. Triviality is not a problem if functions but not modules are domain specific—as domain specificity is no longer an alleged hallmark of modularity but of any computational process. Modules, as architectural subsystems, do not have domains *per se*, only via the (possibly multiple) functions they play.

The rest of their criteria seems compatible with what we said earlier but some of them can be also made less strict.

Let me start with a preliminary definition that will be commented below. The actual notion of mental module that would make the massive modularity interesting, and match the above-mentioned heuristics, is that of a architectural mechanism (for a definition of mechanism see Bechtel, in press); that is:

1. a universally occurring subsystem in a certain species,
2. specified in computational terms,
3. grounded in lower organizational levels,

4. playing a specific role in the system and in behavior, and
5. developing and activating within a certain schedule.

This notion is far more robust than the one explicitly admitted by Carruthers. It could be relaxed by allowing any mental subsystems that do not occur universally, are not innate and develop with individual learning without a fixed schedule. Though subsystems acquired by learning may exist it remains mostly a terminological decision if they should be called “modules”. There is a whole gradation of notions, between the most robust modular notion and the most relaxed version, that is, a simple subsystem that plays a specific role in the system and in behavior. The fourth condition specifies the constraint for functions of modules: only some of the functions of subsystems can qualify as functions of mental modules. These are functions that regulate the system internal organization or interaction with its environment. Functions defined as, say, logical relations or arbitrary disjunctions of system-irrelevant properties cannot qualify as defining roles of mental modules.

Yet, there is another triviality threat even in this version modularity that seems quite robust. Without a clear specification of the outlook of the architecture of the mind—especially of its inter-level interactions—it would imply that any neuron is a mental module. We need a principled way of stipulating proper interactions between the levels of organization. Minds are not simply disorganized modules, they have a highly organized interacting architecture of structured processes.

It is here where the notorious notion of supervenience should be used: the higher-level behaviors must supervene on the lower-level module interactions. If there is no discernible change of behavior (or rather a change of a person-level capacity) when a lower-level process changes, we can suppose that the lower-level process is not a module in our sense. (Note that by ‘discernible’ I mean simply ‘detectable at the proper level.’) That would mean that single neurons in human brains could never become modules in this sense. The modules we are interested in have a certain scale, and we do not want to chunk the neural networks too finely. Any non-trivial notion of a module needs a non-trivial notion of a system, and cognitive systems as we know them are always multi-level. Therefore, we must suppose there are strict inter-level dependencies.

Note that according to this account, a single neuron in the human brain is not a module, while it could be a module in primitive sponges or other simple organisms. So my point is that the module has to be sufficiently sized to contribute to the detectable behavior. Some gradation is possible as long as the notion of behavioral change allows it.

In some situations, even subtle behavioral cues could hint that a module is in operation (for example, a subtle change of behavior of a poker player that bluffs).

Capacities, faculties or modules?

The more robust notion requires further specification of various conditions as it is quite generic and could be turned trivial as well. For it seems that any universal behavioral capacity that is grounded at the neural level is a module, even in the strictest sense (and if physicalism is true, then all capacities are at least partly grounded at the neural level). This objection can be raised against many accounts of functional modularity; the functional subcomponents can be nothing more than capacities or faculties with a fancy name. Functions are always implemented by structures, so what is the advantage of having a structural notion of module if it leads to the same problem?

Take a universal capacity such as the ability to gossip. Gossip is universal across cultures, develops in humans (probably) at certain age, plays a role in behavior, and regulates it. The very first Tooby & Cosmides criterion could be used against this hypothesis but as I already admitted, there are modules with overlapping functions, and they could duplicate (sometimes they must duplicate if they are to compete for resources, cf. Carruthers 2006, 221). So even if there are some modules connected to strategic social knowledge, and there is a curiosity module, we could hypothesize there is still another gossip module. This argument does not seem very strong, but for the sake of discussion, let's agree that gossip specialization could not be easily explained away by other modules (especially when we are quite in the dark about other social modules).

The gossip module has to have a definite computational structure. And this point is of great importance: unless we are able to say what the input of the gossip module is, what are its representations and procedures, and exactly what output it produces, and where the output goes, we are not talking about a module. It is only a capacity. And most personal-level capacities are likely to be implemented by systems of modules, and not by singular, isolated modules. But this much can be admitted by a functional approach of modularity.

If gossip capacity would be realized by a module, there would be a computational mechanism identifiable in the brain, by inspecting its interactions. Without relative isolation, in terms of frequency of interaction at the neural level, a computational mechanism is not a module. But to know the interactions we must know how the module fits

with the rest of the system. Mental architectures hardly seem to be just bunches of self-contained modules with a flat hierarchy (both in time and space), and this makes research harder. This means that the claims about a gossip module—if the module notion is architectural—are non-trivial but at the same time hard to justify.

The defense of the modular account requires that we solve Dennett's Hard Problem (Dennett 2005): What happens next? It is not only about drawing flow diagrams: it has to be shown empirically. Moreover, the proper description of any computational system is a description in terms of the code – and by “code” I mean machine code, a higher programming language or any formal specification of a computational device. A flow-diagram is just a piece of a pseudo-code, so it's just a start. A proper description of state transitions in time is really required to make sense of modularity.

Moreover, the requirement of code specification and lower-level individuation of the computational mechanism cannot be relaxed. It is a basic feature of any computationally specified unit of the architecture of mind. Possible non-universal modules that emerge thanks to the interaction with environment and learning would fulfill it and hence would count as modules.

Of course, the computational code specification requirement is very strong, and currently modular accounts could not possibly fulfill it. However, if we forget about the requirement, the whole modularity hypothesis boils down to a futile *a priori* speculation. Remember that the same behavior could be controlled by many systems with a completely different computational structure. So in order to hypothesize about the internal structure, and to be able to pick the structure that is really implemented, we must try to realistically model the computational architecture of the module. And “realistic modeling” means “modeling as specific as possible”.

Summary

I propose that we need a fairly more robust notion than Carruthers explicitly admits. At the same time, he seems to be using a similarly robust concept in the book. I tried to sketch and improve on that implicit notion—using Simon's architecture of complexity and evolutionary psychology. The notion of modularity is architectural and not simply functional. Its functionality is derived from its organization, and not *vice versa*. Though proponents of modularity explicitly renounce the Fodorian account and usually defend the functional notion (as Carruthers 2006, or

Barrett and Kurzban 2006), some of the problems can be easily solved by using an architectural account. For example, domain specificity of modules does not have to be specified as an individuating property. It is a general property of any computational function and not a special property of modules. A further analysis of how the current notion of function fits the scientific practice is out of the scope of this paper but it seems to fit at least the basic principles of massive modularity defended by Carruthers.

There are several threats to the notion of modularity. First, if it is too broad, the modularity hypothesis becomes trivial, and there is nothing substantial about it. Second, it cannot be fruitfully used without a proper concept of hierarchy and system organization. A concept of module is just a part of ontological dictionary along with such notions as “system”, “function”, and “level.” Third, in the case of mental modules, modularity depends on the controversial claim of computationalism, and computational descriptions are not just flowcharts, they are proper computational specifications. This means we could as well never be able to describe any real human mental module at the proper level of specificity.

The research program fuelled by the massive modularity hypothesis is only about to start. Only if we are able to tackle the questions of the formal specification of the computational structure of mental modules, we can really start the research program of massive modularity. We do not have to start with a full lower-level computational description – such a modularity account would be clearly too *robust* – but providing such a description should be eventually our goal. However challenging this might be, this still seems the best game in town.

References

- Barrett, H. Clark, and Kurzban, Robert. (2006), “Modularity in cognition: Framing the debate.” *Psychological Review*, 113, pp. 628-647.
- Bechtel, Bill. (2007), *Reducing Psychology while Maintaining its Autonomy via Mechanistic Explanation*, in: M. Schouten and H. Looren de Jong (Eds.). *The Matter of the Mind: Philosophical Essays on Psychology, Neuroscience and Reduction*. Oxford, Basil Blackwell.
- Bechtel, W. (in press). *Explanation: Mechanism, modularity, and situated cognition*. In P. Robbins and M. Aydede (Eds.). *Cambridge handbook of situated cognition*. Cambridge, Cambridge University Press.
- Bickhard, Mark. (2000), “Autonomy, Function, and Representation”, *Communication and Cognition - Artificial Intelligence*, vol 17 (3-4), pp. 111-131.

- Callebaut, Werner & Rasskin-Gutman Diego (ed.), (2005), *Modularity. Understanding the Development and Evolution of Natural Complex Systems*, Cambridge, Mass., MIT Press.
- Carruthers, Peter. (2006), *The Architecture of The Mind. Massive Modularity and the Flexibility of Thought*, Oxford, Clarendon Press.
- Collier, John. (2002), *What is Autonomy?* In *Partial Proceedings of CASYS'01: Fifth International Conference on Computing Anticipatory Systems, International Journal of Computing Anticipatory Systems: 12* (2002), pp. 212-221.
- Cosmides, Leda & Tooby, John. (1997), "Dissecting the computational architecture of social inference mechanisms". In *Characterizing human psychological adaptations* (Ciba Foundation Symposium #208), pp. 132-156.
- Cummins, Robert. (1975), "Functional Analysis," *The Journal of Philosophy*, 72, vol. 20, pp. 741-765.
- Dennett, Daniel. (2005), *Sweet Dreams*. Cambridge, Mass., MIT Press.
- Fodor, Jerry. (1983), *Modularity of Mind*. Cambridge, Mass., MIT Press.
- Krohs, Ulrich. (2004), *Eine Theorie biologischer Theorien*, Springer Verlag, Berlin.
- "Functions as Based on a Concept of General Design", *Synthese* (online first, DOI 10.1007/s11229-007-9258-6).
- Millikan, Ruth Garrett. (1984), *Language, Thought, and Other Biological Categories. New Foundations for Realism*. Cambridge, Mass., MIT Press.
- Milner, David & Goodale, Melvyn. (1996), *The Visual Brain in Action*, Oxford, Oxford University Press.
- Prinz, Jesse J. (2006), *Is the Mind Really Modular*. In Robert J. Stainton (ed.), *Contemporary Debates in Cognitive Science*. Blackwell.
- Simon, Herbert, (1996), *The Sciences of the Artificial*. Cambridge, Mass., MIT Press.
- Wimsatt, William. (2000), "Emergence as Non-Aggregativity and the Biases of Reductionism(s)". *Foundations of Science*, 5, 3, pp. 269-297.