The Semantic View of Computation and the Argument from the Cognitive Science Practice

Fabrizio Calzavarini[1,2], Alfredo Paternoster[1,2]

Abstract

According to the semantic view of computation, computations cannot be individuated without invoking semantic properties. A traditional argument for the semantic view is what we shall refer to as the argument from the cognitive science practice. In its general form, this argument rests on the idea that, since cognitive scientists describe computations (in explanations and theories) in semantic terms, computations are individuated semantically. Although commonly invoked in the computational literature, the argument from the cognitive science practice has never been discussed in detail. In this paper, we shall provide a critical reconstruction of this argument and an extensive analysis of its prospects, taking into account some ways of defending it that have never been explored so far. We shall argue that explanatory considerations support at best a weak version of the argument from the cognitive science practice, according to which semantic properties concur with formal syntactic properties in individuating computations in cognitive science, but not a strong version, according to which computation individuation in cognitive science is semantic *as opposed to* formal syntactic.

Keywords: Compuation, Individuation, Cognitive Science, Semantics, Formal syntactic, Neural Computation.

Today computational explanations are ubiquitous across various disciplines, such as cognitive science, neuroscience, and biology. Nevertheless, the issue of how computation is individuated is far from being solved. The traditional view regards computations as identified only by their formal syntactic properties (*formal syntactic view*). According to an alternative view, computations cannot be individuated without invoking semantic properties (*semantic view*). Over the years, the semantic view of computation has been defended on multiple grounds (see Piccinini 2015, pp. 36-44). For instance, it has been said that, since computational devices are individuated by the function they compute, and since functions are individuated in semantic terms (by the ordered couple <domain element, range element>), computational devices are also individuated in semantic terms (e.g., Dietrich 1989; Shagrir 1999). Similarly, it has been argued that a physical device can implement several different automata at the same time, and the contents of the device's states individuate (at least in part) which automaton is relevant for computational individuation. Therefore, computation individuation is (at least) partly semantic. This "master argument" (see Shagrir 2020), which was originally proposed by Oron Shagrir (2001), has provoked a lively debate in recent years (see Piccinini 2008, 2015; Sprevak 2010; Rescorla 2013; Dewhurst 2018; Lee 2018; Cohelo Mollo 2020; Shagrir 2020).

Fabrizio Calzavarini
fabrizio.calzavarini@unibg.it

Alfredo Paternoster
alfredo.paternoster@unibg.it

[1] *Department of Letter, Philosophy, Communication*, University of Bergamo, Italy.
[2] *Center for Logic, Language, and Cognition*, University of Turin, Italy.

Another traditional argument for the semantic view is what we shall refer to as the *argument from the cognitive science practice.* In its general form, this argument rests on the idea that, since cognitive scientists describe computations (in explanations and theories) in semantic terms, computations are individuated semantically. The emergence of this view between the 1980s (e.g., Burge 1986) and the 1990s (e.g., Peacocke 1994) gave rise to a vibrant debate concerning both a descriptive issue (what cognitive scientists *actually do* in their research practice) and a normative issue (what cognitive scientists *should* do), where David Marr's theory of vision was taken as a paradigmatic case-study (e.g., Burge 1986; Davies 1991; Egan 1992, Shapiro 1997). This debate has recently seen a revival, in connection with a renewed discussion on the explanatory practices in cognitive science (e.g., Chalmers 2012; Rescorla 2015a, 2017b). According to Michael Rescorla, for instance, cognitive science explains mental abilities under semantic as opposed to formal syntactic description. For instance, Bayesian cognitive psychology describes the perceptual systems or the motor system as performing probabilistic computations on semantically individuated "hypotheses". In light of this, assuming a "semantically-permeated" notion of computation is explanatorily indispensable since formal syntactic computational descriptions are not relevant to explain core cognitive phenomena.

Although commonly invoked in the computational literature, the argument from the cognitive science practice has never been discussed in detail. In this paper, we shall provide a critical reconstruction of this argument and an extensive analysis of its prospects, taking into account some ways of defending it that have never been explored so far.

We will proceed as follows. After having introduced the debate between the formal syntacticist and the semanticist accounts of computation (§1), we shall discuss what has been taken to be the main objection to the argument from the cognitive science practice, which we shall refer to as the *metaphysical objection*. This is the idea that, since computation individuation is a metaphysical issue concerning the essential properties of computation, it is not affected by explanatory considerations (§2). We shall argue that this objection can be disputed. One might argue that the debate about the individuation of computation ultimately concerns the proper format of computational explanation in cognitive science (§3). If we accept such a view, the argument from the cognitive science practice becomes more credible, and certainly support some formulations of the semantic view according to which semantic properties concur with other non-semantic properties, most notably formal syntactic properties, in individuating cognitive computations. Nevertheless, in our opinion, a careful analysis of the cognitive science practices does not support the stronger claim that semantic properties have explanatory priority over formal syntactic properties, i.e., that cognitive science largely describes (and individuate) cognitive computations in semantic *as opposed to formal syntactic* terms.

1. Computation in Cognitive Science

Arguably the central tenet of classical cognitive science is the idea that cognitive tasks and abilities, such as perception, reasoning, and action programming, are and have to be modeled as computational processes. This is traditionally framed as the idea that mental processes are similar in relevant respect to Turing computations. Scholars who hold classical computationalism have standardly assumed a formal syntactic conception of cognitive computations (e.g., Fodor 1981, Stich 1983, Haugeland 1985, Gallistel and King

2009, Egan 2010). This conception is supposed to account for the fact that brains are not "semantic machines", since all they can do is react differentially to detected (internal) electrochemical differences (Dennett 1981). The brain knows nothing about, say, books or tables; it just processes electrochemical signals. This formal syntactic view is sometimes expressed by saying that one does not need to worry about semantics because, provided that syntax works properly, semantics takes care of itself. The syntactic nature of cognitive computations was somehow taken for granted, basing on the standard method of describing the computational behavior of Turing machines. A Turing machine does, in fact, manipulate a finite set of symbols in accordance with specific rules, and it manipulates them *in virtue of* their formal syntactic properties rather than their semantic ones.

What are formal syntactic properties? As noted by Rescorla (2014b), «[a]t present, we lack a widely accepted analysis of formal mental syntax» (p. 68). In the computational literature, "formal syntactic" can be intended in two different senses, which are rarely distinguished explicitly. According to the narrow sense, 'formal syntactic' has its standard meaning: formal syntactic processes are driven only by rules specifying how complex linguistic structures are constructed by more simple structures. According to the broad sense, which we shall follow in the present paper, "formal syntactic" means merely non-semantic (e.g., Fodor 1978; Rescorla 2017b). The assumption here is that cognitive processes do not "see" the meaning or content of symbols – they do not have access to what symbols stand for. They are only sensitive to formal properties of such symbols.

As Rescorla (2017b) correctly points out, "formal syntactic" in this broad sense also hints to the non-neural character of the relevant properties: they are *multiple realizable* in Putnam's (1967) sense: «physical systems with widely heterogeneous physical properties may satisfy a given syntactic description. Because syntactic description is multiple realizable, it is much more abstract than hardware description» (p. 10). Contemporary defenders of the formal syntactic view, such as Chalmers (2012), tend to agree that cognitive syntax is both non-semantic and multiple realizable.[3] Chalmers glosses syntax in functionalist terms and introduces the concept of *causal topology* and *organizational invariance* to this aim. Causal topology is «the pattern of interaction among parts of the system, abstracted away from the make-up of individual parts and from the way the causal connections are implemented» (2012, p. 337). A property P is organizationally invariant in Chalmers' sense in case «any change to the system that preserves the causal topology preserves P» (2012, p. 337). Accordingly, in its broad sense,

> [t]he syntactic conception […] eschews representation in favour of characterising computation in terms of the abstract functional organisation of the physical system in which it is implemented (O'Brien 2011, p. 387).

---

[3] Note that Rescorla's two conditions for formal syntactic processes (non-semantic and not tied to a particular neural realization) can also be easily satisfied by non-classical and not language-like computational systems such as connectionist or Bayesian models. Therefore, the term "formal syntactic" in this broad sense is supposed to be applicable also to such systems (e.g., Rescorla 2017b; see §4). As observed by an anonymous reviewer, however, one might object that this broad application is unjustified since there is no clear sense in which connectionist or Bayesian models have a syntax (although these systems can indeed be formal). Other terms (e.g., "abstract causal", see §4) might be preferable in this context. Although we agree with this observation, we prefer to keep the "formal syntactic" terminology in order to avoid confusion and preserve consistency with previous literature (especially with Rescorla's writings, which represent the main polemical target of the last sections of the present articles; see §4-5).

Nevertheless, so the story goes, more or less since the mid-1980s (e.g., Burge 1986) some authors started to question this formal syntactic view. Christopher Peacocke (1994; see also 1999), for instance, claimed that computation should have been "content-involving", that is, computational states and algorithms have to be individuated and described in semantic rather than syntactic terms. This is essentially for the reason that, since common-sense mental states are the *explananda* for subpersonal psychology (which is the core of cognitive science), the latter must reflect the (semantic) way intentional psychology individuates mental states. For instance, visual abilities are typically characterized by means of an intentional vocabulary: vision scientists do generally want to account for how our visual system recognizes distal properties (Peacocke 1994). Similarly, behaviors are characterized *relationally*, in terms of the intentions involved and the relevant portion of the world (see Peacocke 1994, p. 307). If one does not assume content-involving computations, she would have a hard time capturing the explananda of scientific psychology:

> [c]ontent-involving computation is peculiarly suited to answering how-questions about content. How does the visual system come to deliver a percept which represents a perceived object as being a certain distance ahead of the perceiving subject? A computational explanation which describes the computation of depth from, say, information about retinal disparity is capable of answering such a how-question. But this too is ruled out by the non-semantic conception of computation, according to which the explaining conditions are also non-semantic. On that conception, the internally individuated explaining conditions cannot magic into existence the complex of non-syntactic relations required for an intentional state to have a certain content (Peacocke 1994, p. 304).

During the 1990s, Marr's theory of vision was used as the main battleground between syntacticists and semanticists. Defenders of the semantic view of cognitive computations, such as Kitcher (1988), Davies (1991), and Shapiro (1993, 1997), all agreed with the early and influential interpretation provided by Burge (1986), according to which visual computations are individuated in Marr's theory by reference to their contents, hence the theory is semantic, or intentional. In Burge's (1986) view, the primary evidence for the semantic character of Marr's theory was the fact that «the top levels of the theory are explicitly formulated in intentional terms» (p. 31). Furthermore, the «intentional primitives of the theory», such as representations of edges, «are individuated by reference to contingently existing physical items or conditions by which they are normally caused and to which they normally apply» (p. 32). According to the then mainstream interpretation (e.g., Shapiro 1993), the attribution of contents in Marr's theory starts at level 1 (the *computational* level), where inputs and outputs of the relevant tasks are formulated in a semantic vocabulary, which mentions distal properties – as is showed by claims such as "the shape of an object is computed from its surface's texture". Computational states and processes at level 2 (the *algorithmic* level) inherit such contents quite naturally.

Opponents of the semantic view, most notably Egan (1992, 1995, 1999), did not negate that, in the informal exposition of the theory, Marr tends to describe visual computations semantically, in terms of features of the distal environment. They also agreed that Marr ascribes environment-specific contents to the computational primitives of the theory where possible. However, according to them, one should not read too much into this fact. First, some of the primitives postulated by Marr correlate only with disjunctive distal properties: «[t]he structures that Marr calls edges sometimes correlate with changes in surface orientation, sometimes with changes in depth, illumination or reflectance» (Egan 1995, p. 195). Other primitives, such as individual zero-crossing, did not correlate with any easily characterized distal properties. More importantly, semantic content does

not play any role for the problem of how to individuate computation. At level 1 of Marr's theory, computational individuation is *mathematical*, or *function-theoretic*, rather than semantic: «[t]he task of a computational mechanism is to compute a certain mathematical function; the initial visual filter, for example, computes the Laplacean convolved with a Gaussian» (Egan 1999, p. 192). Computational states and processes at level 2 might receive semantic content, but this content is not *essential* (more on this later).

In recent years, Marr's symbolic approach has ceased to be the dominant paradigm in the computational science of vision, being in the company with neural networks and Bayesian models (see Chirimuuta 2018). However, the controversy about the role of content in Marr's theory has continued (e.g., Egan 2005, 2019; Shagrir 2010; Ritchie 2018). The custom of appealing to the explanatory practice of cognitive science to support the semantic view of computation has continued too. For instance, in *Origins of Objectivity* Burge (2010) has defended the semantic view of computation again by claiming that current cognitive science is semantic in nature. According to him, «there is no explanatory level in the actual science at which any states are described as purely or primitively syntactical, or purely or primitively formal. One will search textbooks and articles in perceptual psychology in vain to find mention of purely syntactical structures. No explanatory work is given to them» (2010, p. 96). The question of how our visual system recognizes objects, for instance, could not be answered within a computational theory that mentions only formal syntactic properties: «[t]he idea that the visual system is analogous to a purely formal, content-free proof theory does not square with the science» (p. 99). To this explanatory end, so it is claimed, we need content-involving cognitive computations:

> [t]here is no getting around the fact that the laws determining the formation of perceptual states are laws that determine formation of states with representational content. The basic kinds, both explananda and explanans, in perceptual psychology are representational (p.99).

Rescorla has recently reiterated the semantic point with an emphasis on Bayesian models of cognition (e.g., Rescorla 2012, 2015a, 2016, 2017b). In recent years, Bayesian models have been successfully applied to explain several perceptual phenomena, such as color constancy, perceptual illusions, and cue combination, and has been extended to cover sensorimotor processes (see Rescorla 2015a, 2016). The critical point here is that, according to Rescorla, Bayesian cognitive science makes an unavoidable use of semantic vocabulary when characterizing subpersonal computations, thus providing strong support for the semantic view of computation. This is because Bayesian computations typically operate on "hypotheses" that are individuated in representational terms:

> Bayesian models individuate both explananda and explanantia in representational terms. The science explains perceptual states under representational descriptions, and it does so by citing other perceptual states under representational descriptions. For instance, Bayesian models of shape from shading assume prior probabilities over hypotheses about specific distal shapes and about specific lighting directions. The models articulate generalizations describing how retinal input, combined with these priors, causes a revised probability assignment to hypotheses about specific distal shapes, subsequently inducing a unique estimate of a specific distal shape. The generalizations type-identify perceptual states as estimates of specific distal shapes. Similarly, Bayesian models of surface colour perception type-identify perceptual states as estimates of specific surface reflectances. Thus, the science assigns representation a central role within its explanatory generalizations (Rescorla 2015a, p. 702).

According to Rescorla, this fact provides reasonable grounds to claim that computational states and processes in Bayesian psychology are specified in terms of their semantic properties. Critically, Rescorla's points are not limited to Bayesian psychology but are directed towards cognitive neuroscience more in general:

> [t]he critique is not just that representation plays an important role in cognitive science. Fodor and many other formal syntactic computationalists would happily acknowledge that much. The critique is that cognitive science explanation in many areas (including deduction and perception) hinges upon representational content *as opposed* to formal syntax. Scientific practice provides no evidence that the relevant mental processes manipulate formal syntactic items, let alone that the processes are sensitive solely to syntax rather than semantics (2014b, p. 69).

In light of this fact, the kind of formal syntactic computationalism traditionally proposed by Jerry Fodor (1975, 1981), and reiterated by scholars such as Stich (1983), Field (2001), and Chalmers (2012), should be abandoned in favor of content-involving computationalism (see Rescorla 2017a, 2017b). According to content-involving computationalism, cognitive (Turing-style) computations would operate over "semantically permeated" mental symbols, in the sense that the *type-identity* of those symbols would depend on their contents, and hence, so it is claimed, these computations have to be considered content-involving – as Rescorla says, "semantically permeated" (2017a). In Rescorla's view, semantically permeated mental symbols are the natural "ontological" correlates of our pre-theoretical taxonomic scheme for mental entities, which type-identifies mental states, events, or processes partly through their semantic properties (Rescorla 2017a).

2. The Metaphysical Objection

As we have seen, starting from the mid-1990s, several scholars have defended the semantic view of computation, pointing to cognitive science's explanatory practice. In many cases, these scholars' considerations look more like a gesture than a well-defined argument. Despite some differences in their views, we assume that Peacocke, Burge, Rescorla, and the other semanticists would all be disposed to accept this general and straightforward version of the argument from the cognitive science practice:

Argument from the cognitive science practice
(P) Computational descriptions (in explanations and theories) in cognitive science are semantic.
(C) Computations are individuated in semantic terms.

Premise (P) amounts to the claim that, in describing what a given (cognitive) computation does, cognitive scientists make use of a vocabulary that mentions what Rescorla (2017c) calls *veridicality-conditions*. This is an umbrella term designed to cover truth-conditions, accuracy-conditions, fulfillment-conditions, and similar semantic properties. The general idea among semanticists is that semantic description characterizes not just inputs-outputs of computations (lev. 1) but also internal algorithmic states (lev. 2). Despite being apparently framed in a purely descriptive form, premise (P) has both a descriptive and a normative reading, as the *pushmi-pullyu* representations introduced by Ruth Millikan (1995). In the descriptive reading, the claim is equivalent to an empirical generalization over cognitive scientists' explanatory practices – a generalization that

starts from the observation of some representative cases, like Marr's theory or Bayesian psychology. Like any empirical generalization, it allows for exceptions: «[t]here might be *some* areas where cognitive science offers syntactic explanations. For example, certain computational models of low-level insect navigation look both non-neural and non-representational» (Rescorla 2017b, p. 17). The key point is that, *by and large*, cognitive scientists make reference to semantic properties in characterizing cognitive computations.

Premise (P) in its descriptive reading is hardly questionable. Even the fiercest opponents of the semantic view of computation, such as Egan, explicitly recognize that cognitive scientists and neuroscientists make extensive use of representational talks in developing their computational theories and explanations (e.g., Egan 2014, 2019). However, from this fact alone, it does not seem possible to derive a substantial conclusion about computation individuation. What seems to be necessary for establishing the conclusion (C) is, at the very least, an argument for the idea that reference to semantic properties in computational descriptions is justified and not merely a matter of common practice.[4]

The normative reading of (P) gives rise to a more convincing argument but is more controversial. Here the claim is that cognitive scientists *should* refer to semantic properties in computational descriptions if they want to provide adequate theories and explanations. Semantic individuation of computation is supposed to follow from this.

Argument from the cognitive science practice
(P) Computational descriptions (in explanations and theories) in cognitive science should be semantic.
(C) Computations are individuated in semantic terms.

How can the normative reading be defended? As we have seen, a common assumption is that most of the cognitive science's explananda are characterized semantically, either at the pre-theoretical level –that is, at the level of our folk psychological taxonomic scheme (e.g., Rescorla 2017a)– or at a deeper, metaphysical level (e.g., Burge 2010). So, it is claimed, the explanantia of cognitive science, i.e., subpersonal cognitive computations, must also be characterized semantically (e.g., Peacocke 1994; Burge 2010; Rescorla 2017). In other words, sub-personal computational psychology, be it in Marrian or Bayesian form, not only *is* semantic but also *must* be so. As observed by Egan (1995), behind the normative reading of (P) is the intuition that scientific explanations should "match" their explananda, an assumption that has been criticized by several scholars. For instance, Piccinini has argued that the notion that explanantia must be individuated by the same properties that individuate their explananda is at odds with our explanatory practices: «the individuation of explanantia independently of their explananda is an aspect of our explanatory practices. There is no reason to believe that this should fail to obtain in the case of explanations of mental states and processes» (2015, p. 37). The normative reading of (P) is also at odds with a traditional idea that, at least *prima facie*, can be

---

[4] As noted by Coelho Mollo, who takes up a point made by Ramsey (2007), «talk of representation in the cognitive sciences, though widespread, may be largely empty, and we must examine in each case whether it is playing its proper explanatory role» (2020, p. 107). The defenders of the semantic view of computation might object that, without a strong reason to the contrary, we should not assume that scientists are confused or otherwise misguided. Rather, we should take the actual practice in cognitive science as a guide of explanatory success (see Rescorla 2012). However, it is clear that premise (P) in its descriptive reading does not provide strong support for conclusion (C).

attributed to Dennett, which Peacocke has dubbed the *independence claim*. This is the idea that

> […] states involved in subpersonal computations do not have contents that are of the same kind as the contents of personal-level states, nor can they be ascribed contents whose attribution is justified by their power to explain facts about the contents of personal level states (Peacocke 1994, p. 329).

According to Peacocke (1994, pp. 329-33), Dennett's writings contain several arguments in favor of the independence claim. Despite these criticisms, we believe that premise (P) in its normative form remains quite solid. Many scholars have insisted on the indispensability of invoking semantic locutions at the sub-personal level, including scholars that definitely do not belong to the semanticist camp. For instance, Egan has argued that semantic description represents an indispensable "connective tissue" that links the pre-theoretical explanandum and the computational explanans offered in cognitive psychology: «information contained in two-dimensional images», for instance, «is forthcoming only when the states characterized in formal terms by the theory are construed as *representations of distal properties*» (1995, p. 190). Dennett – one of the most authoritative defenders of the "brain as a syntactic machine" analogy – also agrees on this point. He has certainly argued that the *illata* (i.e., posited theoretical entities) of adequate sub-personal psychology will be very different from the kind of personal intentional states posited by folk psychology. This does not mean, however, that these *illata* will not be characterized using intentional labels – that is, as events with content:

> [i]n order to give the *illata* these labels, in order to maintain any intentional interpretation of their operation at all, the theorist must always keep glancing outside the system, to see what normally produces the configuration he is describing, what effects the system's responses normally have on the environment, and what benefit normally accrues to the whole system from this activity. In other words the cognitive psychologist cannot ignore the fact that it is the realization of an intentional system he is studying on pain of abandoning semantic interpretation and hence psychology. […] The alternative of ignoring the external world and its relations to the internal machinery […] is not really psychology at all, but just at best abstract neurophysiology – pure internal syntax with no hope of a semantic interpretation (Dennett 1981, pp. 56-57).

One might argue that the main weakness of the argument from cognitive science lies in the attempt to turn a common explanatory practice in cognitive science, however justified and unavoidable, into a metaphysical doctrine. The current debate on the nature of computation is, in fact, framed on a *metaphysical* sense of "individuation" (see Lowe 2003), according to which what individuates an entity, such a computation, is a property or combinations of properties that that entity possesses *essentially* (e.g., Egan 1995; Piccinini 2015; Lee 2018; Shagrir 2020). It is from this perspective, for instance, that Piccinini argues that «semantic accounts of computation hold that computations are individuated in an essential way by their semantic properties» (p. 33). What counts as an essential property? The issue is rarely discussed in the computational literature. In metaphysics, however, there are two standard answers. According to the *modal* account (e.g., Kripke 1980), an essential property F is a property that an entity *x* possesses always, or *necessarily*; in other words, *x* could not exist if it lacks F. According to the *definitional* account (Fine 1994), an essential property F of an entity *x* is a property that is part of "what *x* is", as elucidated in the definition of *x*. Both definitions of essentiality are (at least tacitly)

assumed in the computational literature. The semantic view so intended is thus the thesis that computation *necessarily* and/or *by definition* involves semantic content.

 Contra Burge, Peacocke, and Rescorla, explanatory considerations from cognitive science, be them in descriptive or normative forms, seem inadequate to support such a metaphysical version of the semantic view of computation. Even if computational descriptions usually operate on entities characterized by means of an intentional, content-involving vocabulary, or even if computational description *should* be content-involving, this does not mean that semantics is essential for computation. This latter claim is a modal and/or conceptual thesis that has to be independently motivated. Thus, even if we accept premise (P) in its strong, normative reading, conclusion (C) does not follow immediately

Argument from the cognitive science practice
(P) Computational descriptions (in explanations and theories) in cognitive science should be semantic.
(C) Computations are *essentially* individuated in semantic terms.

 It seems that there are strong independent modal and conceptual reasons to suppose that semantic properties do not essentially individuate computation. Let us start with modal considerations. As Egan has repeatedly observed in more or less recent years (e.g., Egan 1995, 2010, 2014), one and the same computation can be used to describe different cognitive phenomena in different contexts. Take the computation that is used by Marr to explain the extraction of the primal sketch from the retinal input, i.e., the algorithm that computes the Laplacian of the image convolved with a Gaussian. In Marr's theory, this is a *visual* computation that takes as input changes of intensity at points (x,y) in the retina. In principle, however, the same computation can be used to explain some features of the auditory system, such as the construction of a representation of certain sonic properties from the auditory input. Since, in a different explanatory context, the inputs and outputs (and possibly the computational states) would receive a different semantic interpretation, it can be argued that computations do not necessarily involve representational content. In principle, we can indefinitely vary the context and the semantics while keeping computation (as mathematically described) fixed.[5] Thus, according to Egan (1995),

> [o]nly the mathematical characterization picks out an essential property of a computational mechanism. The intentional characterization is not essential, since in some possible circumstances it would not apply (p. 189) [6]

---

[5] It is important to stress that these considerations are not meant to challenge the premise (P) in its normative force but rather the passage from (P) to the conclusion (C). As we have seen, Egan is explicit in arguing that semantic content is a critical and perhaps necessary component of computational explanations in psychology. She denies, however, the implication that goes from explanatory to metaphysical/individuative considerations. She claims, for instance: «[b]ecause the ascription of distal contents is necessary to explain how a computational process constitutes the exercise of a cognitive capacity in a particular context, I shall call the interpretation that enables the assignment of such distal contents the *cognitive interpretation*. The cognitive interpretation is to be sharply distinguished from the mathematical interpretation […]. Only the latter plays an individuative role» (2012, p. 266). Similarly, she claims that «[r]epresentational content does not play an indivituative role in cognitive theories; but it does play an important explanatory role» (2010, p. 255; see also Egan 1992, 1995, 2014). Of course, this does not mean that, according to Egan, *only* semantic properties have explanatory relevance in cognitive science. In her view, also formal syntactic properties have a critical explanatory role in addition to their individuative role (e.g.,1992, 1995 2014).

[6] Note that these modal considerations seem to hold also in the opposite direction. Consider again the Laplacian example. Assume that another visual system performs edge-detection not by computing the Laplacian (and zero-crossings), but by computing the extreme points of first derivatives (differentiation). In the

Note also that conclusion (C) implies the strong claim that computation individuation is semantic not only in the context of cognitive science but in any field in which computations are invoked, such as artificial computing or robotics. This is because, in the modal reading, «essential means *always*, where "always" refers to any computation, whether actual or possible» (Shagrir 2020, p. 4085). Critically, semantic notions should also be encoded in the definitions of the abstract notions of standard computability theory. This claim appears very controversial Take, for instance, the notion of Turing machine. Turing's computability theory is deeply rooted in the formalist-syntacticist program proposed by Hilbert in the early part of the 20th Century (see O' Brien 2012). As observed by David Chalmers, «[t]he original account of Turing machines by Turing (1936) certainly had no semantic constraints built in. A Turing machine is defined purely in terms of the mechanisms involved, that is, in terms of syntactic patterns and the way they are transformed» (2012, p. 336). Note that the same considerations can be extended to probability theory, which is rooted in Kolmogorov's (1950) formalistic axiomatization, and the Bayesian formalism. As noted by Egan (2020), «[u]nder a natural interpretation, internal structures [of Bayesian models] represent probability distributions. In any event, Bayesian models […] give what I have called a *function-theoretic* characterization; they specify the function, in the mathematical sense, computed by the mechanism. The function is specified intensionally by Bayes' theorem» (pp. 49-50). In this sense, as Rescorla admits (2012), semantic interpretation is not essential for Bayesian models.

3. A Reply (and a Strong Version of the Argument)

As we argued, the main problem of the argument from the cognitive science practice seems to lie in the passage from the premise (P) to the conclusion (C). The problem is that, according to many, the facts that concern computational individuation appear to be largely different and independent from the facts that concern their explanatory role in cognitive science.[7] Importantly, there seem to be independent modal and conceptual reasons to conclude that the notion of computation is non-semantic in its essence, or definition. Granted, these conclusions can be resisted. For instance, it has been said that at least certain computations operate on computational states with intrinsic meaning, so the modal argument, or "argument from reinterpretation", does not work in those cases (Rescorla 2017a, pp. 281-288). Furthermore, it has been said that some abstract computational models are better defined in semantic terms (Rescorla 2012), and that even the Turing Machine formalism is not intrinsically syntactic (Rescorla 2017a, pp. 288-291). This is because such formalism is neutral as to regard the identity of the symbols it operates upon. Therefore, it can also operate on semantically permeated symbols. The debate on such issues is undoubtedly open. The critical point here is that these replies generally do not

---

two cases, the inputs and outputs would receive the same semantic interpretation, whilst the computation (as mathematically described) would be different. Again, this suggests that semantic content is not among the essential properties of computations (we thank an anonymous reviewer for raising this point).

[7] The neat independence between explanatory and metaphysical/individuative considerations seems to be professed not only by defenders of the formal syntactic view of computation, such as Egan, but even by some defenders of the semantic view. For instance, Shagrir (2020) has explicitly argued that the semantic view of computation individuation, being a matter of essential properties, is independent of whether and how semantic properties are used in explanation and theories (e.g., Shagrir 2020, p. 4088).

proceed from explanatory considerations in cognitive science. For this reason, they cannot be easily considered as part of the argument from cognitive science practice.

In this paper, we want to explore a different way in which the argument from cognitive science practice can be defended. First, one might significantly weaken the argument by restricting the scope of its conclusion to individuation *within the context of cognitive science* (and not to computational individuation in general). Second, one might challenge the assumption of neat independence between explanatory and metaphysical/individuative considerations. After all, this assumption is certainly disputable, especially in a debate that belongs to the philosophy of computing and cognitive sciences, and so is primarily motivated by explanatory considerations. This assumption appears to be disputed by some proponents of the semantic view of computation, most notably Rescorla, according to whom «individuation serves explanation» (2017a, p. 286). According to Rescorla, the independence thesis certainly does not apply to the individuation of the cognitive science's explananda, i.e., mental states and processes:

> [w]hen we debate the proper individuation of Mentalese symbols, we are ultimately debating the proper format for psychological explanation. The central issue here is not ontological but explanatory. How do our best psychological explanations type-identify mental states, events, and processes? Should we employ a taxonomic scheme that cites representational properties of mental states? Or should we employ a taxonomic scheme that leaves representational properties underdetermined? (Rescorla 2017a, p. 286).

The critical point here is that, if we accept such conception, the argument from cognitive science practice certainly becomes more credible. According to such conception, individuative considerations are strictly dependent on explanatory issues. As a consequence, the debate about the individuation of computation ultimately concerns the proper format of computational explanation in cognitive science (or in neuroscience, biology, or any other science that invokes computations). It might be said that, since semantic properties appear in any good computational explanation of cognitive phenomena, as we have seen, they belong to the individuative apparatus of computation. Thus, at least in the context of cognitive science, the argument goes, computation individuation is semantic.

Argument from the cognitive science practice
(P) Computational descriptions (in explanations and theories) in cognitive science should be semantic.
(C) Computations *in cognitive science* are individuated in semantic terms.[8]

---

[8] One could argue that conclusion (C) in this form cannot be considered a proper formulation of the semantic view of computation. For it leaves open the possibility that computations in other domains (e.g., artificial computing) are non-semantically individuated. Shagrir (2020) calls "neither semantic nor non-semantic (NSNNS)" the view according to which «computational individuation takes into account semantic properties in some cases but not in others» (p. 4086). Some scholars might dispute this claim. Rescorla, for instance, although initially in line with this characterization (2013), has recently explicitly argued that «[c]ontent-involving computationalists need not say that all computational description is intentional […] They claim only that *some* important computational descriptions are content-involving» (2020). For the sake of argument, we are disposed to assume that it is possible to relativize considerations about computation individuation to specific classes of computational systems (e.g., brains) and to specific scientific fields (e.g., cognitive science). In the context of cognitive science, for instance, the semantic view is the thesis that *cognitive* computations are (or should be) individuated in contentful terms. As we will see, we are interested in ruling out an extreme version of this view, i.e., the idea that semantic *as opposed to formal syntactic* properties should be invoked by computation individuation in cognitive science.

It is critical to observe that that conclusion (C) in itself is still compatible with the idea that other properties are also explanatory relevant in cognitive science on a par with semantic properties, and hence that computation individuation in cognitive science is partly non semantic. This is because it is generally believed that individuative practices in a given scientific domain are not monolithic but depend on multiple factors and serve multiple explanatory functions (e.g., Pemberton 2018). To be sure, this might not be a problem for the defender of the semantic view. According to several formulations of the semantic view, computation individuation takes into account semantic properties; it does not take into account *only* semantic properties. Mark Sprevak (2010), for instance, explicitly argues that «even on [the semantic view], representation would still only be one condition on computational implementation: there are further conditions that a physical computation should satisfy, and additional properties that differentiate physical computations» (p. 112). What kind of non-semantic properties are involved here? Although the issue is rarely discussed in the literature, semanticists usually accept that non-representational neurophysiological properties are also invoked by computation individuation in cognitive science. As noted by Rescorla (2012), the real controversy is whether computational cognitive science needs «an additional level that taxonomizes mental states in formal syntactic terms, without regard to neural or representational properties» (p. 19):

> Everyone agrees that a complete scientific psychology will assign prime importance to neurophysiological description. However, neurophysiological description is distinct from formal syntactic description, because formal syntactic description is supposed to be multiply realizable in the neurophysiological. The issue here is whether scientific psychology should supplement *intentional descriptions* and *neurophysiological descriptions* with *multiply realizable, non-intentional formal syntactic* descriptions (Rescorla 2020).

Critically, conclusion (C) of the argument expressed above –and the argument more generally– is still compatible with a semantic conception of computational individuation in which formal syntactic properties have a prominent explanatory (and thus individuative) role in cognitive science alongside semantic properties. One might say that these two kinds of description occupy distinct levels of explanation. Peacocke seems to endorse such a view (1994; see Rescorla 2014b, 2020). In this sense, we might construe what might be called a *weak* version of the argument from the cognitive science practice:

Weak argument from the cognitive science practice
(P) Computational descriptions in cognitive science should be semantic *in addition to formal syntactic*
(C) Cognitive computations are individuated in semantic terms *in addition to formal syntactic* terms

A few philosophers, however, most notably Burge and Rescorla, seem to endorse a significantly stronger conclusion, which prioritizes semantic description and greatly downgrades the explanatory relevance of formal syntactic description.[9] Such scholars

---

[9] For instance, as we have seen, Burge explicitly argues that no explanatory work is given to formal syntactic properties in actual cognitive science (2010; see §1). In a recent review article, Rescorla describes his own position in a similarly strong form: «Formal syntactic computationalism and content-involving computationalism are compatible […] However, many content-involving computationalists reject formal syntactic computationalism. Tyler Burge and Michael Rescorla question whether formal syntactic description adds any explanatory value to content-involving description. They question whether we can "hive off" a mental state's representational properties to obtain a psychologically significant formal syntactic bearer of those properties» (2014b, p. 69). Similarly strong formulations can be found in several other Rescorla's articles (e.g., 2015a, 2015b, 2016, 2017a, 2017b, 2020). Admittedly, Rescorla sometimes recognizes that

«claim that representational content rather than formal syntax is explanatorily central to numerous core areas, such as the study of perception and deductive reasoning. The basic goal is to delineate computational models that likewise assign explanatory priority to representational content rather than formal syntax» (Rescorla 2014b, p. 1). According to extreme formulations of such view, formal syntactic description has little or no explanatory value in most core areas of cognitive science. Rescorla argues, for instance, that «we can model the mind as a computational system while eschewing any appeal to formal mental syntax. On this view, computational models of the mind can individuate mental states through their representational properties rather than through any alleged formal syntactic properties» (2016, p. 27).[10] Therefore, Rescorla seems to endorse what might be called a *strong* version of the argument from the cognitive science practice:

Strong argument from the cognitive science practice
(P) Computational descriptions in cognitive science should be semantic *as opposed to formal syntactic*.
(C) Cognitive computations are individuated in semantic *as opposed to formal syntactic* terms.

Clearly, in order to evaluate the cogency of this version of the argument, we need to examine the explanatory status of formal syntactic description in cognitive science. In the remainder of this section, we shall reconstruct Rescorla's reasoning in support of premise (P) and conclusion (C) of the "strong" argument, while in the next section, we shall explain why, according to us, such reasoning is not convincing and should be rejected.

As we have seen, form-syntax (in its broad sense; see §1) is standardly characterized in non-semantic and multiple realizable terms. Rescorla is willing to admit that formal syntactic description with these two properties figures prominently in the practice of artificial computing alongside with semantic description.[11] According to Rescorla, in computer science, formal syntactic description is not only standard but also has a pivotal role in the practical purpose of designing and building computer machines. Semantic description of a given function does not specify how to build a machine that computes such function. Syntactic description does it, especially if it is framed in low-level artificial languages (e.g., logic gates or machine descriptions). For «[i]n principle, we know how to build a machine that executes iterated elementary syntactic operations over syntactic items» (2017b, p. 11). Besides, formal syntactic description in artificial computing is

---

there may be some areas in cognitive science in which the formal syntactic scheme is used (e.g., 2017b, p. 17; see §1). Nevertheless, by and large, «current cognitive science does not support any such formal syntactic taxonomic scheme […]. The proposed scheme plays no role within current scientific theories of perception, motor control, mammalian navigation, or numerous other core mental processes. Researchers describe these processes in representational terms» (2021, p. 172).

[10] Note that, in such a view, «there is no rigid demarcation between computational and intentional description. In particular, certain scientifically valuable descriptions of mental activity are both computational and intentional» (Rescorla 2020). Indeed, this does not mean that computationalism adds nothing to representationalism. Instead, it means that computational description (in explanations and theories) of cognitive phenomena are always semantically characterized; in cognitive science, there is no such a thing as semantically indeterminate, multiple realizable computational description (see later).

[11] Granted, Rescorla does not argue that computational individuation in artificial computing is *purely* formal syntactic. According to him, semantic notions have an important explicatory and individuative role in many areas of artificial computing, such as computer science and probabilistic robotics (2017b, pp. 8-9). As we will see, Rescorla's claim is that, differently from cognitive science, artificial computing does assign an important role to formal syntactic description alongside semantic description.

more economical than semantic description since it does not need to specify the «complex causal interactions between the computing machine and its surrounding environment»:

> [b]y focusing solely on "intrinsic" aspects of computation, without seeking to ensure that computational states bear appropriate relations to represented entities, syntactic description carries us much closer to a workable blueprint for a physical system (2017b, p. 12).

Formal syntactic description in artificial computing is also more economical in another sense. According to Rescorla, «[w]hen designing or modifying a computing machine, we often do not care about the exact physical substrate that implements, say, memory registers. We would like a workable blueprint that prescinds from irrelevant hardware details» (2017b, p. 12). Formal syntactic description satisfies this desideratum. Compared to hardware description, therefore, formal syntactic description has the advantage of suppressing low-level implementation (i.e., physical) properties of computation that are not significant for many purposes. In this sense, formal syntactic description in artificial computing

> [h]elps us to manage the enormous complexity of typical computing systems. Designing and modifying complex systems is much easier when we ignore details that do not bear upon our current design goals (Rescorla 2017b, p. 12).

What is true for the practice of computer science does not apply to the study of natural computing systems, such as human minds. According to Rescorla, as we have seen, actual cognitive science, in most areas, does not assign a significant explanatory role to formal syntactic description. For instance, Bayesian models do not cite formal syntactic properties devoid of semantic import. Rescorla (see 2012, 2015a, 2016) noted that it is perfectly possible to imagine a non-semantic Bayesian computational theory in which subjective probabilities are assigned to formal syntactic items rather than semantically individuated hypothesis. Hartry Field (2001, pp. 72, 82, 153-156) has proposed a version of the Bayesian approach along these lines, suggesting that this framework can preserve the semantic approach's explanatory benefits. Nevertheless, the explanatory power of such syntactic version of Bayesian cognitive science is far from being granted:

> [w]e have no reason to believe this conjecture, absent detailed confirmation. Generally speaking, we cannot radically alter how a science individuates its subject matter while preserving the science's explanatory shape. We should not expect that we can transfigure the taxonomic scheme employed by current Bayesian models while retaining the explanatory benefits provided by those models» (Rescorla 2015a, p. 709).

More generally, according to Rescorla, it is always possible to supplement a satisfying semantic and/or neuropsychological theory with a formal syntactic, multiple realizable description. The critical question if we gain any explanatory advantages from this theoretical move. Critics of the semantic view of computation sometimes rely on the superior generality of formal syntactic description of mental phenomena over semantic description (e.g., Chalmers 2012). A formal syntactic description of a given computation has the potential to be applied in other actual and hypothetical situations beyond the specific situation under investigation. In this sense, syntactic description is *scope general*. Nevertheless, according to Rescorla, it is an error to suppose that increased generality is always an explanatory desideratum. First, one can *over*generalize by producing disjunctive or gerrymandered descriptions that add no explanatory import. As Potochnik notes,

«[g]enerality may be of explanatory worth, but explanations can be too general or general in the wrong way» (2010, p. 66). Second, generality does not increase confirmation. For instance, the fact that the Lotka-Volterra equations can model many disparate phenomena other than ecological one does not improve their confirmation value:

> [w]e do not improve ecological explanation by noting that [the Lotka-Volterra equations describe some chemical or economic system when x and y [in these equations] are interpreted as chemical or economic variable […]. What matters for ecological explanation are the ecological interactions described by [the Lotka-Volterra equations], not the causal topology obtained by suppressing ecological variables (2017b, p. 18).

One might ask the reason behind the sharp distinction between artificial computational systems and natural computational systems such as the human mind. Why is formal syntactic description so much more relevant for computer science than cognitive science? According to Rescorla (2017b), this is because the practical advantages of syntactic description are almost entirely irrelevant in cognitive science: «[s]yntactic description enables pragmatically fruitful suppression of representational and hardware properties. No such rationale applies to the scientific study of mental computation. Psychology is not a practical enterprise. Cognitive scientists are not trying to build a computing system. Instead, they seek to explain activity in pre-given computing systems. Constructing an artificial computing system is a very different enterprise than understanding a pre-given computational system» (p. 22). This is why syntactic description has a critical role in the study of artificial computing systems but not for the study of human minds. In light of this fact, the kind of formal syntactic computationalism traditionally proposed by Fodor and reiterated by Chalmers (among other scholars) is on the wrong track:

> I think that these authors distort explanatory practice within actual cognitive science, which evinces no tendency to ground representational description in syntactic description. They also neglect the essentially pragmatic nature of the advantages that syntactic description affords. By heeding the notable differences between artificial and natural computing systems, we may yet articulate more compelling computational theories of mind (Rescorla 2017b, p. 25).

4. The Explanatory Value of Formal syntactic Description

To sum up, we have seen that the argument from cognitive science practice takes on a new light if we restrict the scope of its conclusion and deny the neat independence between metaphysical and explanatory considerations. It might be said that since semantic description of mental computations plays a critical explanatory role in cognitive science, computation in cognitive science should be at least partly individuated in semantic terms. This is certainly plausible. As we argued, the use of semantic vocabulary at the computational level (and possibly at the algorithmic level, too) appears to be necessary to connect a computational theory with its pre-theoretic explananda. The semantic interpretation of cognitive computation also has additional theoretical and explanatory advantages. For instance, it helps the researcher grasp the cognitive system under analysis better and synthetically describe the computation performed by such a system (see, e.g., Egan 2014). We have seen that this version of the argument from cognitive science practice has two possible readings. In its weak reading, the argument is in principle compatible with the idea that formal syntactic description has also a critical role in cognitive science

alongside semantic description. In its strong reading, the argument claims that, since formal syntactic description is explanatory marginal in many core areas of cognitive science, computation individuation in cognitive science is semantic as opposed to formal syntactic. *Prima facie*, Rescorla's considerations against the explanatory value of formal syntactic description in cognitive science might appear convincing. For instance, Rescorla is probably correct in pointing out that scope generality is not always explanatorily valuable in cognitive science (but see later). However, we believe that these considerations ultimately downplay important aspects of the explanatory practices in cognitive science. Thus, the argument from the cognitive science practice is not justified in its "strong" form.

First, Rescorla fails to observe that formal syntactic description (in its broad sense) appears to play an important role in many areas of actual cognitive science, such as computational cognitive neuroscience. In this field, the case of *canonical neural computations* (CNCs) has recently aroused considerable interest (see Carandini & Heeger 2012; Carandini 2012; Chirimuuta 2014; Kaplan 2018). These are defined as «standard computational modules that apply the same fundamental operations in a variety of contexts» (Carandini 2012, p. 5), that is, across different brain areas, different sensory modalities, and even across different species. A prominent example of CNC is divisive normalization, a non-linear operation in which neural responses to external stimuli (e.g., bars in different orientations, or gratings of different contrasts) are «divided by a common factor that typically includes the summed activity of a pool of neurons» (see Carandini & Heeger 2012, p. 51). Originally developed to explain some visual responses in the primary visual cortex, the computational normalization model has been subsequently applied to an impressively wide range of neural phenomena, such as the olfactory system, the high-level visual cortex, the auditory system, or the visual-control system. Other prominent examples of CNCs are linear filtering, recurrent amplification, associative learning, and exponentiation.

CNCs are presented as a «toolbox of computational operations that the brain applies in a number of different sensory modalities and anatomical regions, and which can be described at a higher level of abstraction from their biophysical implementations» (Chirimuuta 2014, p. 58). The exact epistemological status of CNCs and their assimilation to the mechanistic framework have been the subject of intense debate (e.g., Chirimuuta 2014, 2019; Kaplan 2018). The critical point here is that the computational description of cognitive phenomena based on CNCs fits well with the two criteria for syntactic description as individuated by Rescorla. For instance, the normalization model is standardly specified in non-semantic terms, that is, in terms of an equation whose variables can invariably range over a variety of different input profiles (for example, visual stimuli or auditory stimuli). Furthermore, this computational model, as well as many other models involving CNCs, is not supposed to be tied to any particular neural realization:

> [c]rucially, research in neural computation does not need to rest on an understanding of the underlying biophysics. Some computations, such as thresholding, are closely related to underlying biophysical mechanisms. Others, however, such as divisive normalization, are less likely to map one-to-one onto a biophysical circuit. These computations depend on multiple circuits and mechanisms acting in combination, which may vary from region to region and species to species. In this respect, they resemble a set of instructions in a computer language, which does not map uniquely onto a specific set of transistors or serve uniquely the needs of a specific software application (Carandini 2012, p. 507).

According to some scholars, the formal characterization of CNCs does not prevent such computations from having a significant explanatory value in cognitive science. Chirimuuta (2014), for instance, has argued that CNCs are explanatory in that they account

for certain optimality considerations: they explain why a particular cognitive/neural system exhibits a characteristic behavior by appealing to functional utility. Consider the divisive normalization example. The problem, in this case, is to explain why specific neural systems (within the visual cortex, for instance, or the auditory cortex) show non-linear responses to external stimuli. The formal syntactic description of the normalization computation helps in explaining that the non-linear responses enable such systems to transmit more information. In Chirimuuta's view, CNCs explanations are instances of "effective coding explanations", that is, explanations that «ignore biophysical specifics in order to describe the information-processing capacity of a neuron or neuronal population» (2017, p. 851). Critically, such explanations are successful partly because they apply to *classes* to cognitive structures, explaining why such cognitive structures are widespread in the brain. For instance, «[normalization is so widespread] because for many instances of neural processing individual neurons are able to transmit more information if their firing rate is suppressed by the population average firing rate» (Chirimuuta 2014, p. 143).

The case of CNCs is certainly not unique in cognitive neuroscience. Interestingly, *contra* Rescorla, a "formality intuition" appears to have played a role in promoting the so-called *Bayesian Coding Hypothesis*, namely the idea that the «brain represents information probabilistically, by coding and computing with probability density functions or approximations to probability density functions» (Knill & Pouget 2004, p. 713). It is often said that one of the main virtues of this hypothesis is its *unifying power* (see Colombo & Hartman 2015 for discussion), that is, the fact that it is very general and can be applied to many different cognitive tasks and explanatory contexts (e.g., Frinston 2009, 2010; Hohwy 2013). In the neuroscience literature, the Bayesian Coding Hypothesis is often introduced in non-semantic and multiple realizable terms, namely as the hypothesis that a single generic computation is implemented by the brain across a variety of inputs and independently to the biophysical realization (e.g., Ma et al. 2008). As noted by Colombo and Hartman (2015), in the perspective of the Bayesian Coding Hypothesis

> [t]he same type of [probabilistic] relationship can hold not only for visual and tactile information, but for any piece of information associated with a random variable whose distribution has certain mathematical properties. Drawing on this type of relationship, we can derive descriptions of a wide variety of phenomena, regardless of the details of the particular mechanism producing the phenomenon (p. 13).

In this sense, Bayesian computational models can be said to be "abstract causal" representations in Pincock's (2013) taxonomy, namely mathematical, probabilistic representations that abstract away from semantic and causal/physical details (Colombo & Hartman 2015, p. 13). The formality intuition seems to influence Bayesian models not only at computational level, but also at the algorithmic level. As noted by Zednik (2016), defenders of the Bayesian perspective in cognitive neuroscience often invoke a *unification* heuristic, which «highlights those algorithmic-level hypotheses that seem most likely to complement not only the ideal observer model for a single behavioural of cognitive phenomenon, but also the ideal observer models for other phenomena» (p. 21).

Granted, a defender of the semantic view of computation might accept that formal syntactic description is common in cognitive science but still insist, following Rescorla, that formal-syntax is not explanatorily relevant in cognitive neuroscience in spite of its generality and unifying power. Nevertheless, we believe that this theoretical move is an uphill struggle. First off, according to unificationist accounts of explanation, wide scope is a reliable indicator of explanatory depth (e.g., Kitcher 1989). Notably, even scholars

that do not belong to this paradigm have argued that generality/unification might increase confirmation (see Colombo & Hartman 2015, pp. 33-35). For instance, Myrvold (2003) has claimed that, on a Bayesian account, the generality of an explanation contributes to its evidential support for it renders a set of disparate phenomena relevant to each other. This point has been reinforced by the technical discussion provided by Colombo and Hartman (2015, *appendix*), showing that, in a Bayesian framework, coherence with a more general, unified account is a reason to prefer a mechanistic model M_1 over a model M_2. In a similar vein, Pincock (2012) has argued that, in the case of extremely general representations such as abstract, mathematical models, unification increases confirmation because the evidential support can be transferred from one family of models to another.

If the connection between syntactic description and increased confirmation can still be objected, as Rescorla does, the connection between syntax and *fruitfulness* is complicated to deny. Fruitfulness is a diachronic epistemic virtue that is supposed to be linked to explanatory value (see Keas 2018). A theory or explanation is fruitful if it suggests further research that can furnish theoretical insights and new empirical findings. Syntactic description fulfils this desideratum. When specified in non-semantic terms and at a higher level of abstraction from neural implementation, a computational model (e.g., normalization) has the potential to be applied in a variety of cognitive/neural domains, thus stimulating new research questions and new discoveries. Syntactic description can promote what Schurz (2017) calls *analogical* abduction, that is, a type of conceptual abstraction based on an isomorphic or homomorphic mapping between sets of phenomena that were previously considered as unrelated. According to Schurz, «finding an abductive analogy consists in finding the theoretically essential features of the source structure that can be generalized to other domains, and this goes hand-in-hand with forming the corresponding conceptual abstraction» (p. 265). In the computational case, it might be said that syntax provides what is needed to foster analogical extension to multiple cognitive domains.

One might still object that the main advantages of formal syntactic description are not explanatory but pragmatic, and pragmatic considerations have no place in cognitive science. As Rescorla observes, cognitive science is not engineering. This again can be disputed. According to many scholars, engineering reasoning lies at the heart of both classical (see Dennett 1996) and Bayesian cognitive science (see Zednik 2016). As noted by Zednik (2016), «*reverse engineering* strategies begin by developing computational-level models of the phenomena being explained, and then proceed to infer the likely organization of the cognitive system at the algorithmic and implementation levels» (p.18). In Bayesian terms, the first step amounts to provide an ideal observer model of the phenomena to be explained, whilst the second step involves the selection of one algorithm from a space of possible algorithms for computing the ideal observer's behavior. What is critical here is that, as in forward engineering, design and optimality considerations guide the construction of neuro-cognitive theories at these explanatory levels. According to us, the indisputable complexity of the brain and mind makes it pragmatically useful to possess a catalog of abstract, formally-characterized computations and algorithms –like CNCs– that can be used to explain different cognitive phenomena in different contexts.

Based on these considerations, it is clear that formal syntactic description is explanatorily relevant in computational cognitive science, contrary to what Rescorla assumes. Thus, the strong version of the argument from cognitive science practice ultimately fails. Should we conclude that cognitive phenomena are always best studied in formal syntactic terms, following Chalmers, and, consequently, that computational individuation is primarily formal syntactic? In our opinion, the answer to this question is negative. As we

argued, a weak semantic view in which representational and formal syntactic description concur in individuating cognitive computation is preferable. Sometimes cognitive neuroscientists are interested in semantic description, sometimes formal syntactic description is more appropriate, and each choice is guided by explanatory considerations. To the extent that the role of the semantic and the formal syntactic component is balanced in this account, it can also be called a *hybrid* or *pluralistic*. Nevertheless, this perspective should be distinguished from *computational pluralism* recently defended by Jonny Lee (2018). For this latter view applies to computation in general rather than being restricted to individuation in cognitive science. This perspective, however, is consistent with explanatory practices in many domains, where processes are commonly characterized by a rich range of different criteria and types of descriptions (Pemberton 2018).

5. Conclusions

In this paper, we have provided the first critical reconstruction of the argument from the cognitive science practice and the first complete analysis of its prospects. The argument rests on the idea that, since cognitive scientists describe computations (in explanations and theories) in semantic terms, computations are individuated semantically. As we have seen, many scholars believe that explanatory considerations are largely irrelevant to support metaphysical/individuative conclusions about computation in general. According to these scholars, there are independent modal and conceptual reasons to conclude that the notion of computation is non-semantic in its metaphysical essence. As we argued, one might refuse the neat independence of explanatory and metaphysical considerations, and insist that the debate on computation individuation ultimately concerns the proper format of explanation in cognitive science. Even if we accept such a view, we have argued, the prospects of the argument from cognitive science practice ultimately depend on how strong such argument is formulated. In our opinion, at the present state of knowledge, explanatory considerations support at best a weak version of the argument, according to which semantic properties concur with formal syntactic properties in individuating computations in cognitive science, but not a strong version, according to which computation individuation in cognitive science is semantic *as opposed to* formal syntactic. Needless to say, it was not our aim to develop such weak version of the argument from cognitive science practice in full detail. What we aimed to show was just that a careful analysis of the explanatory practice in cognitive science does not license the conclusion that formal syntactic description has little or no explanatory role in actual cognitive science.

Acknowledgments

References

Carandini, M. (2012). From circuits to behavior: A bridge too far? *Nature*, *15(4)*, 507-509.

Carandini. M., & Heeger. D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, *13*, 51-62.

Chalmers, D. (2012). The Varieties of Computation: A Reply, *The Journal of Cognitive Science*, *13*, 213–248.

Chirimuuta, M. (2014). Minimal Models and Canonical Neural Computations: The Distinctness of Computational Explanation in Neuroscience, *Synthese*, *191*,127-153

Chirimuuta, M. (2017). Explanation in Neuroscience: Causal and non-causal. *British Journal of Philosophy of Science*, *9*, 849-880.

Chirimuuta, M. (2018). Vision. In Sprevak Colombo (eds.), *The Routledge Handbook of The Computational Mind*. London: Routledge.

Coelho Mollo, D. (2020). Content Pragmatism Defended. *Topoi*, *39*, 103–113.

Colombo, M., & Hartmann, S. (2015). Bayesian cognitive science, unification, and explanation. *The British Journal for the Philosophy of Science*. doi:10.1093/bjps/axv036.

Burge, T. (1986). *Individualism and Psychology*, *The Philosophical Review*, *95*, 3–45.

Burge, T. (2010). *Origins of Objectivity*. Oxford: Oxford University Press.

Davies, M. (1991). *Individualism and Perceptual Content*, *Mind*, *100*, 461-484.

Dennett, D. (1987). *The Intentional Stance*. Cambridge: MIT Press.

Dennett, D. (1981). Three Kinds of Intentional Psychology, in R. Healey (ed.), *Reduction, Time and Reality*. Cambridge: Cambridge University Press.

Dennett, D. (1994). Cognitive science as reverse engineering: Several meanings of 'top-down' and 'bottomup', in D. Prawitz, B. Skyrms, & D. Westerstahl (Eds.). *Logic, methodology & philosophy of science IX*. Amsterdam: Elsevier Science

Dewhurst, J. (2018). Individuation without representation. *The British Journal for the Philosophy of Science*, *69*, 103–116.

Egan, F. (1991). Must psychology be individualistic? *The Philosophical Review*, *100*, 179–203.

Egan, F. (1992). Individualism, Computation, and Perceptual Content. *Mind*, *101*, 443–459.

Egan, F. (1995). Computation and content. *The Philosophical Review*, *104*, 181–203.

Egan, F. (1999). In Defense of Narrow Mindedness, *Mind and Language*, *14*, 177-194.

Egan, F. (2010). A Modest Role for Content. *Studies in History and Philosophy of Science*, *41*, 253–259.

Egan, F. (2012). Represenationalism, in E. Margolis, R. Samuels, & S. Stich (eds.). *The Oxford Handbook of Philosophy of Cognitive Science*. Oxford: Oxford University Press.

Egan, F. (2014). How to Think About Mental Content, *Philosophical Studies*, *170*, 115–135.

Egan, F. (2019). The Nature and Function of Content in Computational Models, in M. Colombo and M. Sprevak (eds), *The Routledge Handbook of the Computational Mind*. London: Routledge.

Egan, F. (2020). A deflationary account of Mental Representations, in J. Smortchkova, K. Dołrega, T. Schlicht (2020). *What are Mental Representations?* Oxford: Oxford University Press.

Field, H. (2001). *Truth and the Absence of Fact*. Oxford: Oxford University Press.

Fine, K. (1994), Essence and Modality: The Second Philosophical Perspectives Lecture. *Philosophical Perspectives*, *8*, 1–16.

Fodor, J. (1975). *The language of Thought*. New York: Thomas J. Crowell.

Fodor, J. (1981), *Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology*, in J. Fodor, *Representation*, Cambridge: MIT Press.

Fodor, J. (1994). *The Elm and the Expert*. Cambridge: MIT Press.

Friston, K. (2009). The Free-Energy Principle: A Rough Guide to the Brain? *Trends in Cognitive Sciences*, *13*, 293–301.

Friston, K. (2010). The Free-Energy Principle: A Unified Brain Theory? *Nature Review Neuroscience*, *11*, 127–38.

Gallistel, & C.R. King, A. (2009). *Memory and the Computational Brain*. Wiley-Blackwell, Malden

Hohwy, J. (2013). *The Predictive Mind*. Oxford: Oxford University Press

Haugeland, J. (1985). *Artificial Intelligence: The Very Idea*. Cambridge: MIT Press.

Kaplan, D. (2018). Neural Computation, Multiple Realizability, and the Prospects for Mechanistic Explanation, in D. Kaplan (ed.) *Explanation and Integration in Mind and Brain Science*. Oxford: Oxford University Press.

Keas, M. N. (2017). Systematizing the theoretical virtues. *Synthese*, 1-33.

Kitcher, P. (1989). Explanatory Unification and the Causal Structure of the World, in P. Kitcher and W. Salmon (eds), *Scientific Explanation*. Minneapolis: University of Minnesota Press.

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Cognitive Sciences*, *27(12)*, 712–719.

Kripke, S. (1980). *Naming and Necessity*, Cambridge: Harvard University Press.

Kolmogorov, A. (1950) [1933]. *Foundations of the theory of probability*. New York: Chelsea Publishing Company.

Lee, J. (2018). Mechanisms, Wide Functions, and Content: Towards a Computational Pluralism. *The British Journal for the Philosophy of Science*, 2018, axy061.

Lowe, E.J. (2003). Individuation, in Michael J. Loux & Dean W. Zimmerman (eds.). *The Oxford Handbook of Metaphysics*. Oxford: Oxford University Press.

Ma, W. J., Beck, J. M., Latham, P. E. & Pouget, A. (2006). Bayesian Inference with Probabilistic Population Codes. *Nature Neuroscience*, *9*, 1432–8.

Marr, D. (1982). *Vision.* San Francisco: W.H. Freeman.

Millikan, R. (1995). Pushmi-Pullyu Representations. *Philosophical Perspectives*, *9*, 185-200.

Myrvold, W. C. (2003). A Bayesian Account of the Virtue of Unification. *Philosophy of Science*, *70*, 399–23

O'Brien, G. (2012). Defending the semantic conception of computation in cognitive science. *Journal of Cognitive Science*, *12(4)*, 381-399.

Peacocke, C. (1994). *Content, Computation, and Externalism*. Mind and Language, *9*, 303–335.

Peacocke, C. (1999). Computation as Involving Content: A Response to Egan. *Mind and Language*, *14*, 195–202.

Peters, T. (2002), *Timsort description*, accessed June 2015. URL: http://svn.python.org/ projects/python/trunk/Objects/listsort.txt.

Pemberton, J. (2018). Individuating Processes, in O. Bueno, M. B. Fagan, & R. Chen (eds.). *Individuation Across Experimental and Theoretical Sciences*, Oxford: Oxford University Press.

Piccinini, G. (2008). Computation without representation. *Philosophical Studies*, *137*, 205–241.

Piccinini, G. (2015). *Physical computation: A mechanistic account*. Oxford: Oxford University Press.

Potochnik, A. (2010). Levels of explanation reconceived. *Philosophy of Science*, *77(1)*, 59– 72.

Pincock, C. (2012). *Mathematics and Scientific Representation*. New York: Oxford University Press

Putnam, H. (1967). Psychophysical Predicates, in W. Capitan and D. Merrill (eds), *Art, Mind, and Religion*. Pittsburgh: University of Pittsburgh Press.

Putnam, H. (1975). *Mind, language, and reality: Philosophical papers (Vol. 2)*. Cambridge: Cambridge University Press.

Putnam, H. (1988). *Representation and Reality*. Cambridge: MIT Press.

Pylyshyn, Z. (1984). *Computation and Cognition*. Cambridge: MIT Press.

Ramsey, W. (2007). *Representation reconsidered*. Cambridge: Cambridge University Press.

Ramsey, W. (2020). *Defending Representational Realism*, in J. Smortchkova, K. Dołęga, and T. Schlicht, *What are Mental Representations?* Oxford: Oxford University Press.

Rescorla, M. (2009). Cognitive maps and the language of thought. *The British Journal for the Philosophy of Science*, *60*, 377–407.

Rescorla, M. (2012). How to integrate representation into computational modeling, and why we should. *The Journal of Cognitive Science*, *13*, 1–38.

Rescorla, M. (2013a). Against Structuralist theories of computational implementation. *The British Journal for the Philosophy of Science*, *64*, 681–707.

Rescorla, M. (2013b). Millikan on honeybee navigation and communication, in D. Ryder, J. Kingsbury, & K. Williford (Eds.). *Millikan and her critics*. Malden: Wiley-Blackwell.

Rescorla, M. (2014a). The causal relevance of content to computation. *Philosophy and Phenomenological Research*, *88*, 173–208.

Rescorla, M. (2014b). Computational modeling of the mind: What role for mental representation? *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*, 65–73.

Rescorla, M. (2014c). A theory of computational implementation. *Synthese, 191*, 1277–1307.

Rescorla, M. (2015a). Bayesian perceptual psychology, in M. Matthen (Ed.), *The Oxford handbook of the philosophy of perception*. Oxford: Oxford University Press.

Rescorla, M. (2015b). The representational foundations of computation. *Philosophia Mathematica*, *23*, 338–366.

Rescorla, M. (2016). Bayesian sensorimotor psychology. *Mind and Language*, *31*, 3–36.

Rescorla, M. (2017a). From Ockham to Turing—and Back Again, in A. Bokulich & J. Floyd (Eds.), *In Turing 100: Philosophical explorations of the legacy of Alan Turing*. Cham: Springer.

Rescorla (2017b). Levels of Computational Explanation, in T. Powers (ed.). *Philosophy and Computing: Essays in Epistemology, Philosophy of Mind, Logic, and Ethics*. Berlin: Springer.

Rescorla, M. (2017c). Maps in the head?, in K. Andrews & J. Beck (Eds.). *The Routledge handbook of philosophy of animal minds*. Basingstoke: Taylor & Francis.

Rescorla, M. (2020). The computational theory of mind, in E. Zalta (Ed.), (2015, Fall) The Stanford encyclopedia of philosophy. https://plato.stanford.edu/cgibin/encyclopedia/archinfo.cgi?entryDcomputational-mind

Rescorla, M. (2021). "Reifying Representations," *What Are Mental Representations?*, eds. J. Smortchkova, T. Schlicht, and K. Dolega. Oxford: Oxford University Press (2020): pp. 135-177.

Ritchie, B. (2019). The content of Marr's information-processing framework. *Philosophical Psychology*, 32(7), 1078–1099.

Schurz, G. (2017). Patterns of Abductive Inference, in L. Magnani & T. Bertolotti (eds), *Springer Handbook of Model-Based Science*. Berlin: Springer.

Searle, J. (1980). Minds, Brains and Programs. *Behavioral and Brain Sciences*, *3*, 417–57.

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

Shagrir, O. (2010). Marr on Computational-Level Theories. *Philosophy of Science*, *77(4)*, 477-500.

Shagrir, O. (2020). In defense of the semantic view of computation. *Synthese*, *197*, 4083-4108.

Shapiro, L. (1993). Content, Kinds, and Individualism in Marr's Theory of Vision. *The Philosophical Review*, *102(4)*, 489-513.

Shapiro, L. (1997). A Clearer Vision. *Philosophy of Science*, *64*, 131–53.

Sprevak, M. (2010). Computation, Individuation, and the Received View on Representation. *Studies in History and Philosophy of Science*, *41*, 260–270.

Stich, S. (1983). *From Folk Psychology to Cognitive Science*. Cambridge: MIT Press.

Zednik, C., & Jäkel, F. (2018). Bayesian reverse-engineering considered as a research strategy for cognitive science. *Synthese*, *193*, 3951–3985.