# CHAPTER 11

....................................................................................

# REPRESENTATIONALISM

....................................................................................

## FRANCES EGAN

REPRESENTATIONALISM, in its most widely accepted form, is the view that the human mind is an *information-using* system, and that human cognitive capacities are to be understood as representational capacities. This chapter distinguishes several distinct theses that go by the name "representationalism," focusing on the view that is most prevalent in cognitive science. It also discusses some objections to the view and attempts to clarify the role that representational content plays in cognitive models that make use of the notion of "representation."

## 1. SOME REPRESENTATIONALIST THESES

........................................................................................................

The most persuasive argument for representationalism appeals to the fact that evolved creatures such as ourselves behave in ways that are well-suited to achieving their ends given their circumstances. The best explanation of this fact is that they are able to represent both their ends and their circumstances, and that these representational capacities are causally implicated in producing their behavior. More-over, to the extent that an agent's behavioral capacities are flexible, in particular, to the extent that they are sensitive to its changing circumstances, it is plausible to think that its behavior is guided by representational states, at least some of which are themselves a causal product of those circumstances. Sterelny (1990) argues that behavioral flexibility *requires* representation:

> There can be no informational sensitivity without representation. There can be
> no flexible and adaptive response to the world without representation. To learn
> about the world, and to use what we learn to act in new ways, we must be able to

represent the world, our goals and options. Furthermore we must make appropriate inferences from those representations. (21)

Very roughly, the representation of a creature's ends constitutes its *desires*, and the representation of its circumstances its *beliefs*. Representationalism is sometimes understood as a thesis specifically about beliefs and desires. Let us call the view that propositional attitudes—beliefs, desires, fears, intentions, and their kin—are representational states of organisms *Representationalism*$_{PA}$. Representationalism$_{PA}$ is a widely held view, though behaviorists and eliminativists about the attitudes would, of course, deny it.

*Strong Representationalism* is the view that representational mental states have a specific form, in particular, that they are functionally characterizable relations to internal representations. Proponents of Strong Representationalism typically endorse the view that the system of internal representations constitutes a language with a combinatorial syntax and semantics. While strong representationalists typically construe mental representations as language-like, the essential point is that they are *structured* entities over which mental processes are defined. Braddon-Mitchell and Jackson (1996) argue that mental representations may be more analogous to maps than to sentences. Waskan (2006) argues that mental representations are akin to *scale models*. Proponents of strong representationalism include Fodor (1975, 1981, 1987, 2008), Gallistel and King (2009), Pinker (1997, 2005), and Pylyshyn (1984), among many others.

To complete our taxomony of representationalist theses, *Strong Representationalism*$_{PA}$ is the view that propositional attitudes are to be understood as representational states of a specific sort, namely, as functionally characterizable relations to internal representations. According to Strong Representationalism$_{PA}$, to believe that Miles Davis was a genius is to be related in the way characteristic of belief to an internal representation that means *Miles Davis was a genius*.[1] Moreover, proponents of Strong Representationalism$_{PA}$ typically take these internal representations to have a language-like structure.

Several now-classic arguments have been offered in support of Strong Representationalism$_{PA}$. Harman (1972) claimed that logical relations hold among propositional attitudes, and that these relations are essential to their role in predictions and explanations of behavior. If the belief *that snow is white and grass is green* is true, then the belief *that snow is white* is true. In general, if the belief that *p & q* is true, then the belief that *p* is true. Generalizations of this sort presuppose that beliefs have sentential structure. Some beliefs are conjunctions, others conditionals, and so on. Beliefs (as well as desires, fears, and the other propositional attitudes), it is claimed, are part of a language-like system.

Harman's argument trades on the fact that belief *ascriptions* have sentential structure, but it fails to establish that propositional attitudes themselves have logical

[1] The most explicit statement of Strong Representationalism$_{PA}$ is Fodor's *Representational Theory of Mind* (RTM). See Fodor (1975, 1981, 1987, 2008) and Field (1978) for articulation and defense of RTM.

or sentential structure. We ascribe beliefs to subjects using sentences that are conjunctive or conditional, and we make use of the logical relations that hold among the *complements* of belief ascriptions as surrogates to reason about relations that hold among mental states themselves.[2] But it does not follow that the mental states so ascribed are conjunctions or conditionals, or that the relations that hold among these states are of the sort that hold among sentences (or propositions), that is, that they are logical relations. To assume that they are is just to assume what is at issue—that propositional attitudes have a language-like structure. In general, one must be careful not to attribute to thoughts themselves properties of the representational scheme that we use to talk and reason about them.

Fodor (1987) and Fodor and Pylyshyn (1988) argued that certain pervasive features of thought can only be explained by the hypothesis that thought takes place in a linguistic medium. Thought is *productive*: we can think arbitrarily many thoughts. It is also *systematic*: cognitive capacities are systematically related. If a subject can think the thought *John loves Mary*, then she can think the thought *Mary loves John*. The explanation for the productivity and systematicity of thought is that thoughts have a language-like structure. We can think arbitrarily many thoughts for the same reason that we can produce and understand arbitrarily many sentences. Thoughts, like sentences, are composed of a finite base of elements put together in regular ways, according to the rules of a grammar. Systematicity is explained in the same way: systematically related thoughts contain the same basic elements, just arranged differently.

Whether the argument succeeds in establishing that thought is language-like depends on two issues: (1) whether productivity and systematicity are indeed pervasive features of thought; and (2) if they are, whether they can be accounted for without positing an internal linguistic medium.

Thoughts are assumed to be productive in part because they are represented, described, and attributed by public language sentences, a system that is itself productive. But, as noted above, one must be careful not to attribute to thoughts themselves properties of the representational scheme that we use to think and talk about them. It would be a mistake to think that there are substances with infinitely high temperatures just because the scheme we use to measure temperature, the real numbers, is infinite. If thoughts are understood as internal states of subjects that are, typically, effects of external conditions and causes of behavior, then it is not obvious that there are arbitrarily many of them. The size of the set of possible belief-states of human thinkers, like the possible temperatures of substances, is a matter to be settled by empirical investigation.

When we turn to systematicity, the argument falls short of establishing that thought must be language-like. In the first place, it is not clear how pervasive systematicity really is. It is not generally true that if a thinker can entertain a proposition of the form aRb, then he can entertain bRa. One can think the thought *the boy parsed the sentence* but not *the sentence parsed the boy*. Moreover, it is a matter of

---

[2]  See Swoyer (1991) and Matthews (2007) for a detailed argument of this point.

some dispute within the cognitive science community whether connectionist cognitive models, which do not posit a language of thought, might be capable of explaining the systematic relations that do hold among thoughts.[3]

Dennett's 1981 example of a chess-playing computer raises a potentially serious worry for any adherent of Strong Representationalism$_{PA}$. It is plausible to say of the machine that it believes that it should get its queen out early. Ascribing this propositional attitude to the machine allows us to predict and explain its play in a wide range of circumstances. But nowhere in the device's architecture is anything roughly synonymous with "Get the queen out early" explicitly represented, as required by Strong Representationalism$_{PA}$. The consequence of denying that the machine has the appropriate attitude on this ground alone[4] is that doing so would undermine our confidence, more generally, in attitude ascriptions based on the usual behavioral evidence. We have been ascribing propositional attitudes to agents on the basis of their behavior for millennia, without any knowledge of their internal functional or neural architecture. Of course, some of our attitude ascriptions may be false, but their falsity would be revealed, in the typical case, by additional information about the subject's behavioral dispositions—it does not believe it should get its queen out early after all, because when offered the opportunity to do so in a fairly wide range of circumstances, it does something else; she does not believe there is beer in the refrigerator after all, because she is making a special trip to the beer store—not by computational or neural considerations that remain well beyond the ken of ordinary attitude-ascribing folk.[5] Proponents of Strong Representationalism$_{PA}$ often promote it as promising a scientific vindication of our attitude-ascribing folk practices (see, e.g., Fodor 1987). On the contrary, in requiring that every genuine propositional attitude ascribable to a subject corresponds to, or is realized by, an explicit mental representation with the content of that very attitude, Strong Representationalism$_{PA}$ holds those practices hostage to the neural architectural details turning out a particular way, and thus invites an unreasonable eliminativism.

We may conclude that while (regular-strength) Representationalism$_{PA}$ is almost certainly true—propositional attitudes, as they figure in our commonsense predictive and explanatory practices, are conceived of as representational states of organisms—the case for Strong Representationalism$_{PA}$, the view that propositional attitudes are functionally characterizable relations to internal representations that

---

[3] See MacDonald and MacDonald (1996) for the classic papers on this issue, and Cummins (1996), Matthews (1997), and Aizawa (2006) for further discussion.

[4] One might, of course, deny that the chess-playing machine has *any* propositional attitudes, on the grounds that it is not conscious, too cognitively limited, etc. That is another matter.

[5] Certain far-fetched scenarios would also reveal the falsity of ordinary, behavior-based attitude ascriptions. If a subject's behavior was revealed to be caused by remote control by Martian scientists (see Peacocke 1983) then we would withdraw all attitude ascriptions as false. The conditions on correct attitude ascriptions are not *entirely* behavioral. See Egan (1995a) for defense of this view of the attitudes.

have linguistic structure, is inconclusive at best.[6] In any event, let us set Strong Representationalism$_{PA}$ aside for the remainder of this chapter and focus instead on the role that mental representation plays in the cognitive sciences. This strategy is appropriate for several reasons: In the first place, whether or not Strong Representationalism$_{PA}$ is true, cognitive scientists are not committed to its truth, and are not engaged in seeking the "vindication" of our folk practices that its truth would entail. Second, as Von Eckart (1995) notes,

> [W]hen cognitive scientists use the term "mental representation" they are not using it as extensionally equivalent with "propositional attitude". Rather they are using it to refer to computational entities...with representational properties, hypothesized in the context of a scientific research program. (164)

The issue that will concern us in the remainder of this chapter, then, is the role of mental representation in the cognitive sciences.

## 2. REPRESENTATIONALISM IN COGNITIVE SCIENCE—THE STANDARD VIEW

Representationalism, we said, construes the mind as a representational, or *information-using*, device. The notion of a "representational device" is given a precise meaning in cognitive science by Alan Newell's 1980 characterization of a *physical symbol system*.

A *physical symbol system* (hereafter, *PSS*) is a device that manipulates symbols in accordance with the instructions in its program. *Symbols* are objects with a dual character: they are both physically realized and have meaning or semantic content. A realization function $f_R$ maps them to physical state-types of the system. A second mapping, the interpretation function $f_I$, specifies their meaning by pairing them with objects or properties in a particular domain. A given PSS is type-individuated by the two mappings $f_R$ and $f_I$. By this we mean that if either $f_R$ or $f_I$ had been different, the device would be a different (type of) PSS.

The concept of a PSS gives precise meaning to two notions central to mainstream cognitive science: *computation* and *representation*. A *computation* is a sequence of physical state transitions that, under the mappings $f_R$ and $f_I$, executes some specified task. A *representation* is an object whose formal and semantic properties are specified by $f_R$ and $f_I$ respectively.[7]

A PSS, Newell emphasizes, is a universal machine. Given sufficient, but finite, time and memory it is capable of computing any computable function. These

---

[6]  See Stich (1983) and Matthews (2007) for detailed criticism of the RTM.

[7]  Following Fodor's 1980 usage, "formal" here means *non-semantic*.

systems have what Fodor and Pylyshyn (1988) have called a "classical" architecture—an architecture that preserves a principled distinction between the system's representations or data structures and the processes defined over them.

The *physical symbol systems hypothesis* is the idea that the mind is a specific sort of computer, namely a device that manipulates (writes, retrieves, stores, etc.) strings of symbols. The PSS hypothesis is a version of Strong Representationalism, the idea that representational mental states[8]—mental states with semantic content—are functionally characterizable relations to internal representations.

It is not hard to understand the attraction of the PSS hypothesis for philosophers of mind and psychology. Self-styled "computationalists" committed to the view that mental processes are computational processes, notably, proponents of Strong Representationalism$_{PA}$ such as Fodor (1975, 1981, 1987) and Pylyshyn (1984), have hoped that computational models of cognitive processes will eventually mesh with and provide a scientific basis for our commonsense explanatory practices. These practices, as noted above, appeal to content-specific beliefs and desires. For example, it is your belief that there is beer in the refrigerator, together with a content-appropriate desire (to drink a beer, or perhaps just to drink something cold), that explains your going to the kitchen and getting a beer. Appealing to your belief that there is beer at the local bar or your desire to win the lottery fails to provide any explanation of your beer-fetching behavior. Moreover, this behavior is rational just to the extent that it is caused by content-appropriate beliefs and desires. Similarly, according to PSS-inspired computationalism, computational explanations of behavior will appeal to the contents of the symbol strings, or internal representations, the manipulations of which are the causes of our intelligent behavior. But these operations themselves respect what Fodor (1980) has dubbed the *formality condition*—they are sensitive only to *formal* (i.e., *non-semantic*) properties of the representations over which they are defined, not to their content.

The formality condition is often glossed by computationalists as the idea that mental representations have their causal roles in virtue of their syntax. As Pylyshyn (1984) puts the point,

> For every apparent, functionally relevant distinction there is a corresponding syntactic distinction. Thus, any semantic feature that can conceivably affect behavior must be syntactically encoded at the level of a formal symbol structure. By this means we arrange for a system's behavior to be describable as responding to the content of its representations —to what is being represented—in a manner compatible with materialism. (74)

The idea of syntax and semantics marching in lockstep, to produce mechanical reasoning, is of course the fundamental idea underlying theorem proving in logic. But Pylyshyn (1980) elaborates the view as follows:

[8] As opposed to perceptual experiences, pains, and other mental states characterized at least in part by their "qualitative feel." For representational construals of qualitative mental states see Byrne (2001), Dretske (1995, 2003), Lycan (1987, 1996), and Tye (1995, 2000, 2003).

> The formalist view requires that we take the syntactic properties of representations quite literally. It is literally true of a computer that it contains, in some functionally discernible form…what could be referred to as a code or inscription of a symbolic expression, whose formal features mirror (in the sense of bearing a one-to-one correspondence with) semantic characteristics of some represented domain, and which causes the machine to behave in a certain way. (1980, 113)

Pylyshyn's tendency to overlook the fact that a representation has *both* its semantic and syntactic properties only under interpretation—given by the mappings $f_I$ and $f_R$ respectively—leads him to adopt a realist stance toward its syntactic properties that he does not extend to its semantic properties. He says above that the syntactic description of the device is *literally true*, seemingly suggesting that the semantic description of the device is not. What Pylyshyn *should* say is that for any feature of the system—semantic or syntactic—to affect the system's behavior, it must be realized in the device's physical states.[9] The idea that semantics and syntax march in lockstep is, in effect, the idea that both semantics and syntax must be realized in the physical states of the device, semantics in virtue of the dual mappings $f_I$ and $f_R$, and syntax in virtue of $f_R$ alone. Neither semantics nor syntax is an intrinsic feature of the device, although $f_R$ can be seen as specifying the basic causal operations of the device, since it specifies the physical organization relevant for understanding it as a computing device.

Let us focus on the role of so-called "representational content" in computational models of cognitive capacities. Representationalists (of all stripes) tend to endorse the following claims:

(1) The internal states and structures posited in computational theories of cognition are *distally interpreted* in such theories; in other words, the domain of the interpretation function $f_I$ is objects and properties of the external world.

(2) The distal objects and properties that determine the representational content of the posited internal states and structures serve to *type-individuate* a computationally characterized mechanism. In other words, if the states and structures of the device had been assigned *different* distal contents, then it would be a *different* computational mechanism.

(3) The relation between the posited internal states and structures and the distal objects and properties to which they are mapped (by $f_I$)—what we might call the *Representation Relation*—is a substantive, naturalistically specifiable relation. So-called "naturalistic" theories of mental content attempt to specify, in non-intentional and non-semantic terms, a sufficient condition for a mental representation's having a particular meaning. The most popular proposals construe the relation as either information-theoretic (see, e.g., Dretske 1981 and Fodor 1990) or teleological (see

---

[9] See Smith (unpublished) and Marras 1985 for arguments that the formality condition imposes only a realizability constraint on computational theories.

Mill:
vari
grai
less
that
som

We shall call
tion 4 will p
strues both t
"representati
let us turn to

3.

3.1. The (

Strong Repre
processes as
(see his 1966,
tions, memo
chological te
such as Gibs
over the rich
contains mo
held that the
retinal image
ject moves al
are importar
intervening i
to account fc
the richness
tion, propert
higher-order
orientation.
such as being
significant as
could be usec

[10] See the

NITIVE SCIENCE

epresenta-
n some func-
cription of a
aring a one-
sented
980, 113)

*both* its semantic
mappings $f_I$ and
tactic properties
hat the syntactic
nat the semantic
or any feature of
, it must be real-
syntax march in
be realized in the
ngs $f_I$ and $f_R$, and
trinsic feature of
operations of the
erstanding it as a

nt" in computa-
stripes) tend to

al theories of
words, the
operties of the

sentational
to *type-*
n other words,
l *different* distal
anism.
ures and the
y $f_I$)—what we
aturalistically
ntal content
erms, a suffi-
ticular meaning.
r information-
ogical (see

lity condition

Milliken 1984, Dretske 1986, 1995, and Papineau 1987, 1984). While the various proposals on offer face difficulties accounting for the fine-grainedness of mental content, and for misrepresentation,[10] it is nonetheless widely held among philosophers committed to representationalism that the relation between internal structures and their referents satisfies some such naturalistic constraint.

We shall call this package of commitments the *Essential Distal Content View*. Section 4 will present the argument that the Essential Distal Content View misconstrues both the nature of the interpretation function $f_I$ and the role of so-called "representational content" in computational accounts of cognition. First, though, let us turn to some well-known objections to the Standard View.

# 3. CHALLENGES TO THE STANDARD VIEW

## 3.1. The Gibsonian Challenge

Strong Representationalism, as exemplified by the PSS hypothesis, construes mental processes as operations on internal representations. The psychologist J. J. Gibson (see his 1966, 1979 work) held that visual perception is not mediated by representations, memories, concepts, inferences, or any other process characterizable in psychological terms. The difference between so-called "direct theorists" of perception, such as Gibson, and representationalists is often characterized as a disagreement over the richness of the retinal image, with direct theorists arguing that the stimulus contains more information than representationalists are willing to allow. Gibson held that the input to the visual system is not a series of static "time slices" of the retinal image, but rather the smooth transformations of the optic array as the subject moves about the environment, what Gibson (1979) called "retinal flow." There are important constancies in the stimulus that make unnecessary the positing of intervening inferences, calculations, or other processes defined over representations to account for either the subject's ability to perceive size and shape constancies or the richness of visual experience. In addition to intensity and wavelength information, properties directly picked up in the stimulus include, according to Gibson, higher-order properties that remain invariant through movement and changes in orientation. These higher-order invariants specify not only structural properties such as *being a cube* but what Gibson called "affordances," which are functionally significant aspects of the distal scene, such as the fact that an object is edible or could be used for cutting.

---

[10] See the papers in Stich and Warfield (1994) for discussion of some of these worries.

Two fundamental assumptions underlie Gibson's "ecological" approach to visual perception: (1) that functionally significant aspects of the environment structure the ambient light in characteristic ways, and (2) that the organism's visual system has evolved to detect these characteristic structures in the light. Both assumptions are controversial. With respect to (2), representationalists have complained that Gibson provides no account of the mechanism that allegedly detects salient higher-order invariants in the optical array. His claim that the visual system "resonates," like a tuning fork, to these properties is little more than a metaphor.[11] But it should be noted that in claiming that perception of higher-order invariants is direct, Gibson is proposing that the visual mechanism be treated as a black box from the point of view of psychology, because no inferences, calculations, representations, memories, beliefs, or other characteristically *psychological* entities mediate the processing. The *physiological* account of the mechanism's operation will no doubt be very complex. The claim might be plausible if assumption (1) is true—if there is a physically speci-fiable property (or set of properties) of the light corresponding to every perceptible affordance. But for all but the simplest organisms it seems unlikely that the light is structured in accordance with the organism's goals and purposes. More likely, the things that appear to afford eating or cutting or fleeing behavior structure the light in all kinds of different ways. This likelihood has led many perceptual theorists to claim that something like categorization—specifically, the bringing of an object identified initially by its shape, color, or texture under a further concept—is at work when an organism sees an object *as* food, *as* a cutting implement, or *as* a predator. Categorization is then construed as a process defined over representations.

## 3.2. The Challenge from "Embodied" and "Embedded" Cognition

More recently, work in robotics and "artificial life" has inspired a broad-based chal-lenge to the standard view. The MIT researcher Rodney Brooks is the intellectual father of this movement. (See the papers collected in his 1999 work *Cambrian Intel-ligence.*) Brooks's robot Herbert, a self-propelled device equipped with various sen-sors and a moveable arm on top, was designed to motor around the MIT AI lab, collecting empty soda cans and returning them to a central bin. Herbert operates according to what Brooks calls "the subsumption architecture." Its systems decom-pose not into peripheral systems such as vision and motor systems on the one hand, and central systems such as memory on the other, with the latter systems subserving many different tasks, but rather into specialized activity-producing subsystems or skills. Herbert has one subsystem for detecting and avoiding obstacles in its path, another for finding and homing in on distant soda cans, another for putting its hand around nearby soda cans, and so on. What Herbert does not have is a general representation of its environment that subserves all these tasks. Intelligent behavior

emerges from the
construction and 1
following morals f1

We have reachec
esis (H). (C) WI
representations
better to use the
(H) Represen
parts of intellige

Other theorists, fo1
havior and drawn ʒ
that since the mod
there is no reason t
It is a matter
representational.[13]
certainly *informati*
cally differ from th
at least two respec
environment; any 1
within the device,
*ated*; and (2) The
agent's circumstan
not passive structi
they are "action-o1
and suggest an app
is *embodied*.[14] Pro
Thompson, and R
Noe (2004, 2009).[
So far, embodi
An interesting que
own model" will p
tive tasks such as r
sentation of altern
(1991), responding
*inherently* limited:

[12] See, for examp
Port (1995), and Whee
representationalist cla
[13] Clark (1998), f
[14] In these two re
environment and apt
[15] See Anderson
papers in Robbins anc

---

[11] See Fodor and Pylyshyn (1981) for detailed discussion of this criticism.

emerges from the interaction of Herbert's various subsystems, but without the construction and manipulation of representations of its world. Brooks drew the following morals from his work:

> We have reached an unexpected conclusion (C) and have a rather radical hypothesis (H). (C) When we examine very simple level intelligence we find that explicit representations and models of the world simply get in the way. It turns out to be better to use the world as its own model.
>     (H) Representation is the wrong unit of abstraction in building the bulkiest parts of intelligent systems. (1999, 80–81)

Other theorists, following Brooks, have developed similar models of intelligent behavior and drawn similar anti-representationalist conclusions.[12] The general idea is that since the model of some cognitive capacity does not employ representations, there is no reason to think that cognition requires representations.

It is a matter of some dispute whether these models are genuinely non-representational.[13] They count as "representational" in our weak sense—they are certainly *information-using* systems—but such representations as they do have typically differ from those posited in the standard view (Strong Representationalism) in at least two respects: (1) There is no principled demarcation between device and environment; any representations posited are not conceived of as contained entirely within the device, hence the idea that cognition, so modeled, is *embedded* or *situated*; and (2) There is no strict demarcation between the representation of the agent's circumstances and its goals, but more significantly, its representations are not passive structures over which computational operations are defined. Rather, they are "action-oriented," as Clark (1998) puts it—they *both* describe a situation and suggest an appropriate behavioral response to it—hence the idea that cognition is *embodied*.[14] Proponents of embedded or embodied cognition include Varela, Thompson, and Rosch (1991), Clancy (1997), Clark (1998), Gallagher (2005), and Noe (2004, 2009).[15]

So far, embodied/embedded devices are capable only of fairly simple behaviors. An interesting question is how fruitful Brooks's strategy of "using the world as its own model" will prove to be in modeling and understanding more complex cognitive tasks such as reasoning and problem solving, which seem to require the representation of alternative courses of action and counterfactual circumstances. Kirsh (1991), responding to early work, suggests that the embedded/situated approach is *inherently* limited:

---

[12] See, for example, Beer (1995), Harvey, Husbands, and Cliff (1994), van Gelder and Port (1995), and Wheeler (1996). See Chemero (2000) for a critical discussion of these anti-representationalist claims.

[13] Clark (1998), for example, thinks that they are; Gallagher (2005, 2008) thinks they are not.

[14] In these two respects, they are suggestive of Gibson's *affordances*, which are located in the environment and apt for appropriate behavior.

[15] See Anderson (2003) for an excellent review of the embodied cognition literature, and the papers in Robbins and Aydede (2008) for views for and against situated/embedded cognition.

> Situationally determined activity has a real chance of success only if there are enough egocentrically perceptible cues available. There must be sufficient local constraint in the environment to determine actions that have no irreversibly bad downstream effects. Only then will it be unnecessary for the creature to represent alternative courses of actions to determine which ones lead to dead ends, traps, loops, or idle wandering. From this it follows that if a task requires knowledge about the world that must be obtained by reasoning or by recall, rather than by perception, it cannot be classified as situation determined. (171)

Clark (1998) also expresses skepticism that embedded/embodied accounts will be able to adequately model and explain higher cognitive processes, and suggests that a hybrid approach, incorporating both embedded/embodied processes and structured representations of the sort advocated by the standard view, is most promising. The hard question then is how the two schemes might be coordinated, given their disparate commitments about the fundamental nature of cognition.

## 3.3. The Chomskian Challenge

Noam Chomsky has argued in his recent work that the so-called "representational" states invoked in accounts of our cognitive capacities are not genuinely representational and that they are not correctly construed as about some represented objects or entities. Discussing computational vision theory he says,

> There is no meaningful question about the "content" of the internal representations of a person seeing a cube under the conditions of the experiments ... or about the content of a frog's "representation of" a fly or of a moving dot in the standard experimental studies of frog vision. No notion like "content", or "representation of", figures within the theory, so there are no answers to be given as to their nature. The same is true when Marr writes that he is studying vision as "a mapping from one representation to another..." (Marr, 1982, p. 31)— where "representation" is not to be understood relationally, as "representation of". (1995, 52–53)

The idea that "representation" should, in certain contexts, not be understood relationally as in "representation of x," but rather as specifying a monadic property, as in "x-type representation," can be traced to Goodman (1968).[16] So understood, the

---

[16] According to Goodman,

Saying that a picture represents a so-and-so is thus highly ambiguous as between saying what the picture denotes and saying what kind of picture it is. Some confusion can be avoided if in the latter case we speak rather of ... a "Pickwick-picture" or "unicorn-picture" or "man-picture". Obviously a picture cannot, barring equivocation, both represent Pickwick and represent nothing. But a picture may be of a certain kind—be a Pickwick-picture or a man-picture—without representing anything." (1968, 22)

Goodman claims that the locution "representation of Pickwick" is syntactically ambiguous. On one reading it has the logical form of a one-place "fused" predicate—"Pickwick-representation"—where "Pickwick" is, in Quine's 1960 terminology, "syncategorematic." Chomsky is not committed to this syntactic thesis.

individuating condition of a given internal structure is not its relation to an "intentional object," there being no such thing according to Chomsky, but rather its role in cognitive processing. Reference to what looks to be an intentional object is simply a convenient way of type-identifying structures with the same role in computational processing.

The point applies as well to the study of the processes underlying linguistic capacities:

> [H]ere too we need not ponder what is represented, seeking some objective construction from sounds or things. The representations are postulated mental entities, to be understood in the manner of a mental image of a rotating cube, whether the consequence of tachistoscopic presentations or of a real rotating cube or of stimulation of the retina in some other way, or imagined, for that matter. Accessed by performance systems, the internal representations of language enter into interpretation, thought, and action, but there is no reason to seek any other relation to the world. (Chomsky 1995, 53)

Chomsky rejects the idea that intentional attribution—the positing of a domain of objects or properties to which internal structures stand in a *meaning* or *reference* relation—plays any explanatory role in cognitive science. Intentional construals of David Marr's 1982 theory of vision, such as Burge (1986), Chomsky claims, are simply a misreading, based on conflating the theory proper with its informal presentation. As Chomsky puts it, "The theory itself has no place for the [intentional] concepts that enter into the informal presentation, intended for general motivation" (1995, 55).

Chomsky himself has not spelled the argument out explicitly, though the motivation for his recent anti-representationalism is not hard to find.[17] As theories of our perceptual and linguistic capacities have become increasingly removed from commonsense, it becomes quite forced to say that the subject *knows* or *believes*, say, the *rigidity assumption* (Ullman 1979) or the *minimal link condition* (Chomsky 1995). Chomsky (1975) was willing to say that subjects "cognize" the principles posited in cognitive theories, but these contents—*that objects are rigid in translation* or *that derivations with shorter links are preferred over derivations with longer links*—do not look much like the sorts of things that subjects could plausibly be said to know, believe, etc. They are not inferentially promiscuous, not accessible to consciousness, and so on.

It is particularly unclear what independent objects the structures posited in accounts of our linguistic capacities represent. Among the candidates are elements of the public language,[18] elements of the speaker's idiolect, or, as Georges Rey (2003a, 2003b, 2005) has recently suggested, linguistic entities such as nouns, verb phrases,

---

[17] Though see Collins (2007) for the view that Chomsky has always been an anti-representationalist.

[18] Chomsky himself is skeptical of the notion of a "shared public language." See the papers in Chomsky (2000).

phonemes, and so on—what Rey calls "standard linguistic entities" (SLEs). SLEs, Rey argues, are to be understood as "intentional inexistents," objects of thought, akin to such fictional entities as Zeus or Hamlet, that do not exist. Discussion of the merits and demerits of these various proposals is beyond the scope of the present chapter. Chomsky, for his part, rejects them all, insisting that talk of represented objects is intended simply for informal exposition and plays no genuine role in the theory.

# 4. Rethinking the Standard View

Chomsky is, in effect, an *eliminativist* about representational content. He denies that the internal structures posited in computational theories are distally interpreted as representations of external objects and properties (Claim 1 of the Essential Distal Content View), and hence that computational mechanisms are type-individuated by a domain of external objects and properties (Claim 2).[19] Any reference to such a domain in computational accounts, he claims, is merely "informal presentation, intended for general motivation."

This section shall spell out a view of representation in computational cognitive theories according to which Chomsky is correct in denying the Essential Distal Content View, but nonetheless wrong in denying to representational content a genuine explanatory role. Chomsky's view fails to make clear the role played by the interpretation function $f_I$ in computational accounts, and leaves mysterious how representational content could aid in the "informal motivation" of a computational theory. These points will be illustrated by reference to two computational models from different cognitive domains.[20]

David Marr's well-known explanatory hierarchy distinguishes three distinct levels at which a computational account of a cognitive capacity is articulated. Disputes about whether computational theories type-individuate the mechanisms they characterize by their representational content turn on how the level of description that Marr called the *theory of the computation* should be interpreted. The theory of the computation provides a *canonical description* of the function(s) computed by

---

[19] It is consistent with Chomsky's stated views that there *is* a substantive, naturalistically specifiable relation between posited structures and distal objects and properties (Claim 3 of the Essential Distal Content View), though Chomsky himself would regard the idea that theories of mind and language must respect such a "naturalistic constraint" as a manifestation of "methodological dualism," the idea that the study of language and mind, unlike scientific inquiry in other domains (which is allowed to be self-policing), should be held to independent, "philosophical" standards. See Chomsky (1994).

[20] See Egan (1995b, 1999, 2003) for defense of this account of the role of content in David Marr's theory, in particular.

the mechanism.
mean the chara
tion or taxonom
$f_I$. The canonica
important point
tationally chara
amples illustrate
Marr (1982)
the initial filteri
describe what th
portant charact
description: the
337). As it happ
image and calcu
characterization
intended for ger
input values rep
*intensity*. The cc
well-understooc
with the transdu
The second
*representation a*
cuitry or neura
early visual proc
tions, and disc
them. It is at this
tions and proce:
etc.) and the pr(
$f_R$, which (ideal
states and proce
The second
and Wise's 200;
involving object
fixated ahead. T
(t) at Xt. The pr(
an infinite num|
at Xt. But for m
one of these tra;
locity. Shadmel
accomplished.
The overall
first problem is:
*nematics* involv(
from proprioce|

the mechanism. It specifies what the device does. By a "canonical description," we mean the characterization that is decisive for settling questions of type-individuation or taxonomy. The canonical description is given by the interpretation function $f_i$. The canonical description is therefore a semantic characterization. But this is the important point: the canonical description of the function computed by a computationally characterized mechanism is a *mathematical* description. A couple of examples illustrate the point.

Marr (1982) describes a component of early visual processing responsible for the initial filtering of the retinal image. Although there are many ways to informally describe what the filter does, Marr is careful to point out that the theoretically important characterization, from a computational point of view, is a mathematical description: the device computes the Laplacean convolved with the Gaussian (1982, 337). As it happens, it takes as input light intensity values at points in the retinal image and calculates the rate of change of intensity over the image. But this *distal* characterization of the task is, as Chomsky might put it, an "informal" description, intended for general motivation. *Qua* computational device, it does not matter that input values represent *light intensities* and output values the rate of change of *light intensity*. The computational theory characterizes the visual filter as a member of a well-understood class of mathematical devices that have nothing essentially to do with the transduction of light.

The second and third levels of Marr's explanatory hierarchy describe a *representation and algorithm* for computing the specified functions, and the circuitry or neural hardware that implement the computations. Marr's account of early visual processing posits primitive symbol types—edges, bars, blobs, terminations, and discontinuities—and selection and grouping processes defined over them. It is at this second level that the theory posits symbol structures or representations and processes defined over them. These symbol structures (edges, bars, blobs, etc.) and the processes that operate on them are type-individuated by the mapping $f_R$, which (ideally, when fully specified) characterizes them at the level of physical states and processes, independent of the cognitive capacities that they subserve.

The second example, from an entirely different cognitive domain, is Shadmehr and Wise's 2005 computational theory of motor control. Consider a simple task involving object manipulation (see Figure 1). A subject is seated at a table with eyes fixated ahead. The hand or *end effector* (ee) is located at Xee, and the target object (t) at Xt. The problem is simply how to move the hand to grasp the object. There are an infinite number of trajectories from the hand's starting location Xee to the target at Xt. But for most reaching and pointing movements, the hand moves along just one of these trajectories; it typically moves along a straight path with a smooth velocity. Shadmehr and Wise (2005) describe one way in which the task might be accomplished.

The overall problem can be broken down into a number of subproblems. The first problem is: how does the brain compute the location of the hand? *Forward kinematics* involves computing the location of the hand (Xee) in visual coordinates from proprioceptive information received from the arm, neck, and eye muscles, and
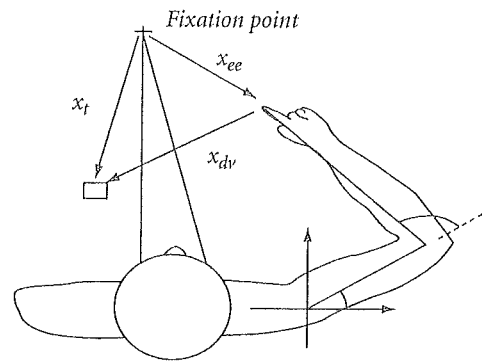
Figure 1. A motor control task.

information about the angles of the shoulder and elbow joints. Informally, this process coordinates the way the hand looks to the subject with the way it feels. The brain also has to compute the location of the target ($X_t$), using retinal information and information about eye and head orientation.

The second problem, computing a plan of movement, involves computing the *difference vector*, that is, the displacement of the hand from its current location to the target's location. But this "high level" plan specifies a displacement of the hand in visual coordinates. This visually oriented plan has to be transformed into a specification of the joint rotations and muscle forces required to effect the displacement. So, the third problem, involving the computation of *inverse kinematics and dynamics*, is how the high level motor plan, corresponding to a difference vector, is transformed into joint angle changes and force commands. Reaching and pointing movements involve continuous monitoring of target and hand location, with the goal of reducing the difference vector to zero. There are a number of complicating factors. For example, incidental eye and head movements require continuous updating of the situation. Deceleration of the hand should be smooth, to avoid knocking over the target.

Summarizing, the account decomposes the overall task into three computations, and specifies the function computed in each in precise mathematical terms:

(1) $f(\theta) = X_{ee}$, *forward kinematics*, the computation of hand location, in eye-centered coordinates, from propriocentric information and information about joint angles;

(2) $X_t - X_{ee} = X_{dv}$, *the difference vector*, the difference between the target location and initial hand position in eye-centered coordinates; and

(3) $f(X_{dv}) = \Delta\theta$, *inverse kinematics*, the computation from the high-level movement plan, in eye-centered coordinates to a required change of joint angles.

The motor control mechanism characterized by Shadmehr and Wise is not a physical symbol system; its operations are not interpreted in the account as manipulations of symbols. Nor does the account of the mechanism's implementation

decompose neatly into representation and algorithm (Marr's level 2) and neural realization (Marr's level 3). Rather, the three computations that constitute the motor control mechanism are characterized as analog processes and realized in neural networks in the posterior parietal cortex, the premotor cortex, and the primary motor cortex respectively. The details need not concern us here.

The important point is that in both examples, the canonical description of the task executed by the device, the function(s) computed, is a mathematical description. As noted above, this description characterizes the mechanism as a member of a well-understood class of mathematical devices. A crucial feature of this characterization is that it is "environment neutral": the task is characterized in terms that prescind from the environment in which the mechanism is normally deployed. The mechanism described by Marr computes the Laplacean of the Gaussian whether it is part of a visual system or an auditory system, in other words, independently of the environment—even the *internal* environment—in which it is normally embedded. In fact, it is not implausible to suppose that each sensory modality has one of these same computational mechanisms, since it just computes a curve-smoothing function. The same point holds for the motor control mechanism characterized by Shadmehr and Wise. A mariner who knew the distance and bearing from his home port to his present location and the distance and bearing from his home port to a buried treasure could perform the same computation to compute the course from his present location to the treasure. In both cases, it is the abstract mathematical description that type-individuates the mechanism or process, not what Chomsky would call the "informal" description that characterizes the mechanism as computing *changes of light intensities* or the *displacement between target and hand location.*

To summarize: The characterization of a computational process or mechanism made available by the interpretation function $f_I$ – the mapping that provides a canonical description of the function computed by the mechanism, and hence (along with the realization function $f_R$) serves to type-individuate it—is an abstract mathematical description. This semantic interpretation does not provide a *distal* interpretation of the posited internal states and structures; the specified domain is not external objects and properties in the subject's environment, but rather mathematical objects. The upshot is that the Essential Distal Content View mischaracterizes the semantic interpretation of a device given by the interpretation function $f_I$. Its domain is *not*, as Claim 1 holds, external objects and properties. And the representation relation determined by the mapping $f_I$ does not, as Claim 3 holds, satisfy a constraint of the sort that proponents of a "naturalistic" semantics for mental representation have been hoping for. The interpretation maps the posited internal states and structures to a domain of *abstracta*, rather than external objects or properties required for an information-theoretic or teleological relation.

If this account is correct, then what should we make of the idea that visual states represent such visible distal properties as *depth* and *surface orientation*, and motor control states represent *hand location* and *shoulder angle*? Are such distal contents

explanatorily idle, as Chomsky claims? And if they aid in "general motivation," how precisely do they do that?

Ordinary, distal representational contents do not serve to type-individuate a computational mechanism, as Claim 2 of the Essential Distant Content View holds, but they do serve several important explanatory functions. The questions that antecedently define a psychological theory's domain are usually couched in intentional terms. For example, we want a theory of vision to tell us, among other things, how the visual system can detect three-dimensional distal structure from information contained in two-dimensional images. A characterization of the postulated computational processes in terms of distal objects and properties enables the theory to answer these questions. This characterization tells us that states of the system co-vary, in the normal environment, with changes in depth and surface orientation. It is only under an interpretation of some of the states of the system as representations of depth and surface orientation that the processes given an environment-neutral, mathematical characterization by a computational theory are revealed as *vision*. Thus, content ascription plays a crucial *explanatory* role: it is necessary to explain how the operation of a mathematically characterized process constitutes the exercise of a cognitive capacity in the environment in which the process is normally deployed. The device would compute the same mathematical function in any environment, but only in some environments would its doing so enable the organism to see.

This is the most important function of representational content. Because the ascription of distal contents is necessary to explain how a computational process constitutes the exercise of a cognitive capacity in a particular context, we 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 specified by $f_I$. Only the latter plays an individuative role.

To recap: When the computational characterization is accompanied by an appropriate cognitive interpretation, in terms of distal objects and properties, we can see how a mechanism that computes a certain mathematical function can, in a particular context, subserve a cognitive function such as vision or reaching and pointing. So when the input states of the Marrian filter are described as representing *light intensities* and the output states *changes of light intensity* over the image, we can see how this mechanism enables the subject to detect significant boundaries in the scene. When the input states of the mechanism that computes inverse kinematics are described as representing *displacement in visual space* and the output states *changes in joint angles*, we can see the role that the mechanism plays in the subject's successfully grasping the target object.

The account presented here draws a sharp distinction between the computational theory proper—the mathematical description made available by the mapping $f_I$, which (together with $f_R$) type-individuates the mechanism—and the distal characterization, made available by the cognitive interpretation that accompanies it and explains the contribution of the abstractly characterized mechanism to the

GNITIVE SCIENCE

motivation," how

ype-individuate a
ntent View holds,
questions that an-
ouched in inten-
ong other things,
ure from informa-
of the postulated
erties enables the
that states of the
depth and surface
ates of the system
rocesses given an
putational theory
lanatory role: it is
racterized process
nent in which the
me mathematical
vould its doing so

ntent. Because the
putational process
ntext, we shall call
tents the cognitive
istinguished from
plays an individu-

npanied by an ap-
properties, we can
tion can, in a par-
aching and point-
representing light
image, we can see
boundaries in the
nverse kinematics
the output states
ays in the subject's

een the computa-
lable by the map-
m—and the distal
hat accompanies it
mechanism to the

larger cognitive life of the organism. We can also understand how representational content, specified by the cognitive interpretation, while not type-individuating a computational mechanism, can, as Chomsky puts it, provide "general motivation" for the theory.

The cognitive characterization is essentially a *gloss* on the more precise account of the mechanism provided by the computational theory. It forms a bridge between the abstract, mathematical characterization that constitutes the explanatory core of the theory and the intentionally characterized pretheoretic explananda that define the theory's cognitive domain. When the processes given a precise mathematical specification in the theory are construed, under interpretation, as representations of such distal properties as *edges*, or *joint angles*, the account can address the questions that motivated the search for a computational theory in the first place, such as *how are we able to see the three-dimensional structure of the scene from two dimensional images?*, or *how are we able to move our hand to grasp an object in sight?*

To call the cognitive characterization a "gloss" is not to suggest that the ascription of representational content is unprincipled. The posited states and structures are not interpretable as representations of distal visible properties (as, say, *object boundaries*, or *depth* or *surface orientation*) unless they co-vary with tokenings of these properties in the subject's immediate environment. It would be a mistake, though, to conclude that the structures posited in computational vision theories must (even in the gloss) represent their normal distal cause, and to find in these theories support for a *causal* or *information-theoretic* theory of content.[21] Some structures—zero-crossings in Marr's account, for example—are interpreted as representations of proximal features, in particular, as *discontinuities in the image*. The ascription of content is sometimes driven by purely expository considerations, such as allowing us to keep track of what the process is doing at points in the processing where the theory posits structures that do not correlate neatly with a salient distal property tokening. Even within a single cognitive interpretation, no single, privileged relation is assumed to hold between posited structures and the elements to which they are mapped. The choice of a cognitive gloss is governed by explanatory considerations, which we can, following Chomsky, characterize as "informal motivation."

An implication of the foregoing account of the role of representational content in computational models is that cognitive science has no need for a naturalistic semantics—the specification of non-intentional and non-semantic sufficient conditions for a mental state's having the meaning it does. Whatever the *philosophical* interest of such an account, it would hold little interest for the computational cognitive theorist. Content ascription in computational models is motivated primarily by explanatory considerations that a naturalistic semantics does not address.

---

[21] See Dretske (1981) and Fodor (1990) for examples of information-theoretic accounts of content.

## 5. Concluding Remarks on Strong Representationalism

*Representationalism*, recall, is the view that the human mind is an information-using system. So understood, representationalism is hard to deny. As Sterelny (1990) noted, it is hard to see how our various cognitive capacities could be explained except by positing states that are both sensitive to the world around us and causally involved in producing behavior.

*Strong Representationalism* goes further, construing human cognitive states as relations to internal representations, thereby positing a sharp distinction—inherent in the notion of a physical symbol system—between data structures or representations on the one hand, and the processes defined over them on the other. This view certainly has its attractions, not the least of which is that it purports to explain *how* thinkers can be information-using systems. They use information by manipulating representations of that information. This idea requires a distinction between the part of the system that uses representations and the representations themselves, which is exactly what the data structure/process distinction enforces.

A real attraction of cognitive models that conform to Strong Representationalism is their *explanatory transparency*. Symbols are structures ready-made for semantic interpretation—they just *are* objects with both formal and semantic properties, characterized by $f_R$ and $f_I$ respectively. Symbol structures—representations—are, in effect, "hooks" on which a semantic interpretation can be hung. Moreover, the information in physical symbol systems is *accessible for use*, encoded in exactly the features of the structures to which computational processes are sensitive. Thus, physical symbol systems are said to *explicitly represent* the information that they encode. And we can track the flow of information in the system by keeping track of the operations on the encoding structures.[22],[23] This is an important explanatory role played by what we call the "cognitive interpretation."

The structure/process distinction inherent in Strong Representationalism is not inevitable. There are ways to eschew the distinction while preserving the central idea of (regular strength) representationalism—that the mind is an *information-using device*. The Shadmehr and Wise motor control mechanism described above is an example. It is not a physical symbol system; its computations are characterized as analog processes and realized in neural networks. Parallel distributed processing (PDP) systems, analog relaxation systems, massive cellular automata, and other kinds of computational mechanisms for which the structure/process distinction is not preserved are *not* ready-made for interpretation. Often, in PDP systems, no distinct state or part of the network serves to represent any particular object,

[22] This is the point of positing what Egan and Matthews (2006) call "intentional internals."
[23] Though see Kirsch (1990) for an argument that even PSS models may not be as transparent as commonly thought.

property, or prop
many units, conne
of any given objec
network. It become
that our paradigm
spatially compact,
terpretation). The
cannot tell by look
which of its spatio
difficult to track th
tation is the same
spicuous as *cogniti*

It is even possi
might be sensitive
the world without
tational job. If hur
then understandin

The explanato
distinction explair
human cognitive
relations to intern
Strong Representa

Aizawa, K. (2006). *Tl*
Anderson, M. (2003)
Beer, R. (1995). Com
  and Van Gelder
Braddon-Mitchell, D
  Blackwell.
Brooks, R. (1999). *Ca*
Burge, T. (1986). Indi
Byrne, A. (2001). Inte
Chemero, A. (2000).
  *Science* 67: 625–
Chomsky, N. (1975).
——. (1994). Natu
  *Journal of Philos*

[24] Ramsey (2007)
*representational.*

property, or proposition. Rather, the encoding of information is distributed over many units, connection strengths, and biases, with the result that the representation of any given object, property, or proposition is widely scattered throughout the network. It becomes quite forced to talk of "representations" in such systems,[24] given that our paradigm of a representation is the printed word, a discrete object that is spatially compact, movable, and, most important, *meaningful* (but only under interpretation). These systems are far from explanatorily transparent. One often cannot tell by looking at the computational and engineering details of the system which of its spatiotemporal parts are candidates for interpretation. It can be quite difficult to track the flow of information in these systems. But the point of interpretation is the same as for PSS systems—to make the computational processes perspicuous as *cognitive* processes.

It is even possible that mental representation is a *global* affair. A whole system might be sensitive to environmental changes and hence be involved in representing the world without any localizable part of the system doing any particular representational job. If human cognitive capacities are the result of processes of this sort, then understanding and modeling these capacities will be very challenging.

The explanatory transparency of systems that respect the structure/process distinction explains the attraction of Strong Representationalism, the view that human cognitive processes are to be understood as functionally characterizable relations to internal representations. While this may be a reason to hope that Strong Representationalism is true, it is not a reason to believe it.

# REFERENCES

Aizawa, K. (2006). *The Systematicity Arguments*. Dordrecht, The Netherlands: Kluwer.
Anderson, M. (2003). Embodied cognition: A field guide. *Artificial Intelligence* 149: 91–130.
Beer, R. (1995). Computational and Dynamical Languages for Autonomous Agents. In Port and Van Gelder (eds.), *Mind as Motion*. Cambridge, MA: MIT Press.
Braddon-Mitchell, D., and Jackson, F. (1996). *Philosophy of Mind and Cognition*. Oxford: Blackwell.
Brooks, R. (1999). *Cambrian Intelligence*. Cambridge, MA: MIT Press.
Burge, T. (1986). Individualism and psychology. *The Philosophical Review* 95: 3–45.
Byrne, A. (2001). Intentionalism defended. *The Philosophical Review* 110: 199–240.
Chemero, A. (2000). Anti-representationalism and the dynamical stance. *Philosophy of Science* 67: 625–47.
Chomsky, N. (1975). *Reflections on Language*. New York: Pantheon Books.
——. (1994). Naturalism and dualism in the study of language and mind. *International Journal of Philosophical Studies* 2: 181–209. (Reprinted in Chomsky 2000.)

[24] Ramsey (2007) argues that it is not appropriate to characterize such systems as *representational*.

————. (1995). Language and nature. *Mind* 104: 1–61.

————. (2000). *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.

Clancy, W. (1997). *Situated Cognition: On Human Knowledge and Computer Representation*. Cambridge: Cambridge University Press.

Clark, A. (1998). *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press.

Collins, J. (2007). Meta-scientific eliminativism: A reconsideration of Chomsky's review of Skinner's *Verbal Behavior*. *British Journal for the Philosophy of Science* 58: 625–58.

Cummins, R. (1996). Systematicity. *Journal of Philosophy* 93: 591–614.

Dennett, D. (1981). A Cure for the Common Code?, In *Brainstorms*. Cambridge, MA: MIT Press.

Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press.

————. (1986). Misrepresentation. In R. Bogdan (ed.), *Belief: Form, Content and Function*. New York: Oxford University Press.

————. (1995). *Naturalizing the Mind*. Cambridge, MA: MIT Press.

————. (2003). Experience as representation. In E. Villanueva (ed.), *Philosophical Issues*, 13: 67–82.

Egan, F. (1995a). Folk psychology and cognitive architecture. *Philosophy of Science* 62: 179–96.

————. (1995b). Computation and content. *The Philosophical Review* 104: 181–203.

————. (1999). In defense of narrow mindedness. *Mind and Language* 14: 177–94.

————. (2003). Naturalistic Inquiry: Where Does Mental Representation Fit In?. In L. Antony and N. Hornstein (eds.), *Chomsky and his Critics*, Oxford: Blackwell.

Egan, F., and Matthews, R. (2006). Doing cognitive neuroscience: A third way. *Synthese* 153: 377–91.

Field, H. (1978). Mental representation. In *Erkenntnis*, 13: 9–61.

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

————. (1980). Methodological solipsism considered as a research strategy in cognitive science. *Behavioral and Brain Sciences* 3: 63–73.

————. (1981). *RePresentations: Philosophical Essays on the Foundations of Cognitive Science*. Cambridge, MA: MIT Press.

————. (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press.

————. (1990). *A Theory of Content and Other Essays*. Cambridge, MA: MIT Press.

————. (2008). *LOT 2: The Language of Thought Revisited*. Oxford: Oxford University Press.

Fodor, J., and Pylyshyn, Z. (1981). How direct is visual perception? Some reflections on Gibson's "ecological approach." *Cognition* 9: 139–96.

————. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition* 28: 3–71.

Gallagher, S. (2005). *How the Body Shapes the Mind*. New York: Oxford University Press.

————. (2008). Are minimal representations still representations? *International Journal of Philosophical Studies* 16: 351–69.

Gallistel, C. R., and King, A. P. (2009). *Memory and the Computational Brain*. Oxford: Wiley-Blackwell.

Gibson, J. (1966). *The Senses Considered as Perceptual Systems*. Oxford: Houghton Mifflin.

————. (1979). *The Ecological Approach to Visual Perception*. Oxford: Houghton Mifflin.

Goodman, N. (1968). *Languages of Art*. Indianapolis: Bobbs-Merrill.

Harman, G. (1972). *Thought*. Princeton, NJ: Princeton University Press.

Harvey, I., Husbands, P., and Cliff, D. (1994). Seeing the Light: Artificial Evolution, Real Vision. In D. Cliff (ed.), *From Animals to Animats 3*. Cambridge, MA: MIT Press.

Kirsh, D. (1990). When Is Information Explicitly Represented? In P. Hanson (ed.), *Information, Content, and Meaning*, Vancouver: University of British Columbia Press.

———. (1991). Today the earwig, tomorrow man. *Artificial Intelligence* 47: 161–84.

Lycan, W. (1987). *Consciousness*. Cambridge, MA: MIT Press.

———. (1996). *Consciousness and Experience*. Cambridge, MA: MIT Press.

Marr, D. (1982). *Vision*. New York: Freeman.

Marras, A. (1985). The Churchlands on methodological solipsism and computational psychology. *Philosophy of Science* 52: 295–09.

Matthews, R. (1997). Can connectionism explain systematicity? *Mind and Language* 12: 154–77.

———. (2007). *The Measure of Mind: Propositional Attitudes and Their Attribution*. Oxford: Oxford University Press.

Macdonald, C., and Macdonald, G. (1996), *Connectionism: Debates on Psychological Explanation*. Oxford: Blackwell.

Millikan, R. (1984). *Language, Thought, and Other Biological Categories*. Cambridge, MA: MIT Press.

Newell, A. (1980). Physical Symbol Systems. *Cognitive Science* 4: 135–83.

Noe, A. (2004). *Action in Perception*. Cambridge, MA: MIT Press.

———. (2009). *Out of Our Heads*. New York: Hill and Wang.

Papineau, D. (1987). *Reality and Representation*. Oxford: Blackwell.

———. (1993). *Philosophical Naturalism*. Oxford: Blackwell.

Peacocke, C. (1983). *Sense and Content*. Oxford: Oxford University Press.

Pinker, S. (1997). *How the Mind Works*. New York: Norton.

———. (2005). So how does the mind work? *Mind and Language* 20: 1–24.

Pylyshyn, Z. (1980). Computation and cognition: Issues in the foundation of cognitive science. *The Behavioral and Brain Sciences* 3: 111–32.

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

Quine, W. (1960). *Word and Object*. Cambridge, MA: MIT Press.

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

Rey, G. (2003a). Chomsky, Intentionality, and a CRTT. In L. Antony and N. Hornstein (eds.), *Chomsky and His Critics*. Oxford: Blackwell.

———. (2003b). Intentional Content and a Chomskian Linguistics. In A. Barber (ed.), *Epistemology of Language*. Oxford: Oxford University Press.

———. (2005). Mind, intentionality, and inexistence: An overview of my work. In *The Croatian Journal of Philosophy* 5: 389–415.

Robbins, P., and Aydede, M. (2008). *The Cambridge Handbook of Situated Cognition*. Cambridge: Cambridge University Press.

Shadmehr, R., and Wise, S. (2005). *The Computational Neurobiology of Reaching and Pointing: A Foundation for Motor Learning*. Cambridge, MA: MIT Press.

Smith, B. (unpublished). Semantic attribution and the formality condition.

Sterelny, K. (1990). *The Representational Theory of Mind*. Oxford: Blackwell.

Stich, S. (1983). *From Folk Psychology to Cognitive Science: The Case against Belief*. Cambridge, MA: MIT Press.

Stich, S., and Warfield, T. A. (1994). *Mental Representation: A Reader*. Oxford: Blackwell.

Swoyer, C. (1991). Structural representation and surrogative reasoning. *Synthese* 87: 449–508.

Tye, M. (1995). *Ten Problems of Consciousness*. Cambridge, MA: MIT Press.
———. (2000). *Consciousness, Color, and Content*. Cambridge, MA: MIT Press.
———. (2003). Blurry Images, Double Vision, and Other oddities: New Problems for Representationalism. In Q. Smith and A. Jokic (eds.), *Consciousness: New Philosophical Perspectives*, Oxford: Clarendon Press.
Ullman, S. (1979). *The Interpretation of Visual Motion*. Cambridge, MA: MIT Press.
van Gelder, T., and Port, R. (1995). It's about time. In R. Port and T. van Gelder (eds.), *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge, MA: MIT Press, 1–44.
Varela, F., Thompson, E., and Rosch, E. (1991). *The Embodied Mind*. Cambridge, MA: MIT Press.
Von Eckart, B. (1995). *What is Cognitive Science?* Cambridge, MA: MIT Press.
Waskan, J. (2006). *Models and Cognition*. Cambridge, MA: MIT Press.
Wheeler, M. (1996). From Robots to Rothko. In M. Boden (ed.), *The Philosophy of Artificial Life*. Oxford: Oxford University Press.

WHAT i
tion is
technol
While 1
restrict
issue o
accordi
stood a
much 1

Sec
and we
been at
tion? T
will foc

Fir
and 2 c
to the 1
in man
*is* anytl
unders
unders
connec