Synthese (2009) 166:547–591
DOI 10.1007/s11229-008-9375-x

# The interactivist model

**Mark H. Bickhard**

Received: 16 December 2007 / Accepted: 11 July 2008 / Published online: 30 July 2008


**Abstract** A shift from a metaphysical framework of substance to one of process enables an integrated account of the emergence of normative phenomena. I show how substance assumptions block genuine ontological emergence, especially the emergence of normativity, and how a process framework permits a thermodynamic-based account of normative emergence. The focus is on two foundational forms of normativity, that of normative function and of representation as emergent in a particular kind of function. This process model of representation, called interactivism, compels changes in many related domains. The discussion ends with brief attention to three domains in which changes are induced by the representational model: perception, learning, and language.

**Keywords** Interactivism · Substance · Process · Emergence · Normativity · Hume · Kim · Normative function · Representation · Millikan · Dretske · Fodor · Cummins · Perception · Learning · Evolutionary epistemology · Language · Social ontology

The interactivist model of representation and cognition is an action and interaction based approach—it is roughly Pragmatic in that sense. It involves fundamentally different assumptions about representation than those made in standard models in the literature, and, more deeply, a fundamentally different metaphysical framework from the substance, structure, and particle frameworks that are still dominant in most of philosophy, cognitive science, and psychology. It is based on a process metaphysical framework—also roughly Pragmatic.

M. H. Bickhard (✉)
Lehigh University, 17 Memorial Drive E., Bethlehem, PA 18015, USA
e-mail: mark@bickhard.name
URL: http://www.bickhard.ws/

The interactivist model began with a relatively narrow, primarily cognitive, focus, but the interaction and process assumptions involved in that early narrower model differed significantly from dominant assumptions in related domains. Consequently, it proved impossible to integrate the early model with theories in those related domains with which it had to interface—e.g., as an account of interactive representation must interface with an account of language.

It became necessary to either abandon the original model, or to extend it into interrelated domains. This scenario of overcoming incompatible modeling assumptions by extending the core model has, since the original confrontation with theories of language, been replayed many times with regard to a wider and wider range of phenomena, and with respect to deeper and deeper levels of theoretical and metaphysical assumptions involved. Over the course of some years of such growth, it has become a skeletally systematic philosophy and theory, addressing multiple phenomena and layers of metaphysical assumptions across a wide range of topics. I will focus here primarily on normative biological function, representation, and cognitive issues.

The discussion proceeds in three parts: (1) an outline of an underlying metaphysical stance within which the remainder of the article is developed, (2) a model of the emergence of normativity, in the forms of normative function and representation, and (3) a partial discussion of some further implications for phenomena such as perception and language.

## 1 Metaphysics: process versus substance

The heritage of the Greeks includes some fundamentally aporetic metaphysical positions that still dominate contemporary thought. Among other consequences, phenomena such as the emergence of normativity are rendered conceptually impossible. Before turning to some more recent manifestations, I briefly look at the origins of this conceptual tradition.

Against Heraclitean flux, Parmenides argued that change is not even possible: For A to change into B would require that A disappear into nothingness and B emerge out of nothingness. Nothingness cannot exist, therefore change is not possible.

Parmenides' argument against nothingness turned on the Greek notion that speaking or thinking about something was akin to pointing to that something (Campbell 1992). Because nothingness cannot be pointed to, it cannot exist, and, therefore, cannot participate in change. Change, therefore, cannot occur. This is an early example of the difficulties that nothingness and falsity have posed throughout Western thought: consider, for example, the difficulties that recent thinkers from Russell to Fodor have had with the possibility of representing something that does not exist, or with false representation.

Certainly the Greeks took these problems seriously, and attempted to provide solutions or dissolutions of them. In particular, Empedocles proposed that everything is constituted out of four substances—earth, air, fire, and water—and that such basic substances did *not* change, thus satisfying the Parmenidean constraint against nothingness and emergence. Within this framework, the appearance of change and differences could be accounted for in terms of varying mixtures of the substances and changes in those mixtures (Guthrie 1965; Wright 1997). Similarly, Democritus proposed non-divisible Parmenidean wholes—atoms—which did not change, but that apparent change could be accounted for in terms of various and changing configurations of such atoms (Guthrie 1965; Taylor 1997). Aristotle's version of the four elements of earth, air, fire, and water, was much more sophisticated than that of Empedocles, and change was possible at this level in Aristotle's framework, but, nevertheless, there remained an unchanging, therefore Parmenides-satisfying, base of prime matter (Gill 1989).[1]

The acceptance of such a metaphysical ground of unchangingness, be it substance or particle, has strong and potentially problematic consequences:

1. The explanatory default is lack of change—inertness—and change requires explanation.
2. Emergence is impossible. Empedoclean substances can mix, but there is no way for a fifth substance to emerge.
3. The metaphysical realm of substances or particles and their properties and interrelations, characterized in terms of cause and fact, is split off from the realm of mental kinds of phenomena, properties, and interrelations—of intentionality, normativity, and modality.

Within a metaphysical framework that incorporates this split between substance and mind, there are three general possibilities. The two realms of substance and mentality can be accepted as a basic metaphysical division, and accounts of the world attempted in terms of the two realms. Aristotle's matter and form, Descartes' two kinds of substances, Kant's two realms of noumena and phenomena, and early analytic philosophy's diremption between the factual world of science and that of social and linguistic normativity, are all examples.

Another possibility would be to attempt to account for the world in terms of just the "mental" realm, yielding an idealism like that of Hegel, Green, or Bradley. Or one might try to explain it in terms of just the "substance" or material realm alone, like Hobbes, Hume (on many interpretations), Quine, and most contemporary philosophers and psychologists.

It might seem intuitively attractive to try to integrate phenomena across the two realms via some sort of emergence of one out of the other, but emergence is

[1] The word substance is used both for substance as stuff and for substance as entity or particular thing. Substance as particular thing is especially common in traditions derived from translations of Aristotle. In such frameworks, stuff is generally referred to as 'matter'. I am using 'substance' to refer to some kind of stuff—an unchanging substratum—out of which the world is constituted, thus including 'matter' in most of its uses. Issues of unity, with which Aristotle was rightly concerned (Gill 1989), having to do with the unity that makes an object or entity different from an aggregate, do not have the same form in a process framework. Such issues do, however, at least partly overlap with issues of process stability, with which I will be centrally concerned.

Substance as particular thing is the most common usage in ontology, but substance as stuff is also found (e.g., Butchvarov 1999; Belsey 1995; Graham 1997, 2006; Guthrie 1965; Reale 1987; Trusted 1999). Substance as stuff is more common in scientific usage—e.g., gold or water as substances (Robinson 2004). The confusion and potential for confusion here is sufficient that some have elected to translate Aristotle's notion of substance as particular thing as "entity" rather than as "substance" (e.g., Campbell 1992; Owens 1978).

precisely one of the conceptual possibilities that the substance/particle metaphysics was intended to preclude.[2]

### 1.1 Hume

One powerful statement of and argument for such a split is Hume's argument against the possibility of deriving norms from facts. Hume does not actually elaborate the details of such an argument (Hume 1978), but the standard interpretation offers the following considerations (Schurz 1997): If the premises of a valid argument contain only factual terms, then it follows, so it is claimed, that any conclusion can only contain factual terms. Any terms not already available in the premises must be introduced by definitions based on terms in the premises or on those previously defined. Consequently, any terms in the conclusion that are not in the premises can be back-translated through their definitions—in each case substituting the defining phrase or clause for the defined term—until the conclusion is stated solely using terms that are in the premises. Since those are by assumption strictly factual, then the conclusion must similarly be strictly factual. This argument has had a formative impact on the Western tradition.

Note that, although the argument is stated in terms of facts and norms, "is" and "ought", its general form is that nothing new can be obtained in the conclusion beyond configurations of what is already available in the premises. In this general form, the argument precludes any kind of emergence, and is a manifestation of the restriction to combinatorics of the basic substance/particle framework.[3]

This argument, however, is unsound. It is based on the assumption that all definition permits back-translation, and that is false. In particular, implicit definition does not.

Implicit definition was recognized and introduced most especially in the late 19th century by Hilbert in his axiomatization of geometry (Hilbert 1971; Kneale and Kneale 1986; Otero 1970). The general idea of implicit definition is that the terms involved are defined by their interrelationships—in contemporary model theoretic terms, an axiom system implicitly defines its class of models (Chang and Keisler 1990; Shapiro 2005a,b). The notion also applies to definitions of single terms as locations in relationships with other already available terms, as well as to non-formal senses (Hale and Wright 2000).[4]

Implicit definition has been largely ignored in the contemporary mainstream of analytic metaphysics, at least until very recent times, for two reasons: (1) many implicit definitions permit satisfaction by multiple and non-isomorphic models, thus not satisfying a criterion of unique differentiation, and (2) Beth's theorem states that implicit definition and explicit definition are equally powerful, thus apparently giving no reason to bother with implicit definition (Doyle 1985). Point (1) is not relevant to my current

[2] Clearly, this would have to be some kind of ontological emergence. An epistemologically based emergence would simply leave all of the metaphysical issues untouched.

[3] Note that the issue here is not forms of inference, but forms of definition and of the kinds of representations that those definitions permit.

[4] For generalizations of implicit definition within a dynamic framework, see Bickhard and Terveen (1995) and Bickhard (in preparation).

purposes (and holds for defining, e.g., natural numbers, primarily when restricted to first order logic: Shapiro 1991, 1997, 2005a,b),[5] and point (2) holds in first order predicate logic with infinite models, but does not hold in all combinations of kinds of logics and kinds of models. In general, the equivalence of implicit and explicit definition is only extensional even when it does hold, and in many other combinations of logics and models, implicit definition is more powerful (Kolaitis 1990; Quine 1966). Implicit definition cannot successfully be dismissed.

Most importantly for current purposes, implicit definition does not permit back-translation. There is no defining phrase or clause that can be substituted for a defined term. Even in formal cases, it is a matter of mathematical and logical discovery to determine what characterizes the class of models for a given implicit definition.

But, if implicit definition is a legitimate form of definition, and if implicit definition does not permit back-translation, then Hume's argument is based on a false premise. Hume's argument is unsound, and this barrier, at least, to the possibility of an emergence account of normativity and mentality is removed.

### 1.2 Kim

A second barrier to the possibility of emergence that I will consider is an argument by Jaegwon Kim that all causality resides in the lowest level of physical entities, whatever that may be. In this view, any new causal properties that might be manifest in higher level organization will be mere causal regularities, regularities of the causal interactions of the particles in that particular configuration (Kim 1991). In this view, organization is just the framework, the stage, on which all genuine causal interactions take place.

The primary tool in Kim's arsenal is the pre-emption argument: Consider a supposed higher level emergent, supervenient on an organization of lower level elements, and ultimately on the basic particles of physics. Any causal consequences of the higher level emergent will either be redundant with respect to the causal consequences of the lower level particle interactions, and therefore preempted by the causality of those lower level interactions—in which case the higher level is not itself causally efficacious—or else the higher level introduces some additional causal power—in which case the micro-level of physics is not causally closed. Because the causal closure of the micro-physical world seems to be a premise of both scientific and philosophical naturalism, the conclusion is that no emergent can have causal efficacy. All "emergent" causality is epiphenomenal relative to micro-causality.[6]

[5] Issues about implicit definitions requiring isomorphic models have primarily to do with definitions of number systems. There is no even prima facie relevance of this issue for implicit definitions of, for example, a mathematical group or field.

[6] This reasoning depends on the organizational *relations* among the lower level constituents being included in the supervenience base. If those relations were not included, then the higher level causal regularities could *not* be accounted for in terms of that supervenience base. The argument, then, applies against dualisms (precluded by the causal closer assumption) and assumptions of new causal laws coming into application with new organizations, as with some of the classical British emergentists (McLaughlin 1992).

More recently, Kim has explored possibilities opened up by removing relations from the definitions of various kinds of base (e.g., micro-level base, supervenience base, and so on) and by considering notions

This is a powerful argument. But it is not sound, and the false assumption that is involved points toward a different metaphysical framework within which to explore issues of emergence. In particular, the descent of causal power downward through various levels of organization that this argument yields constitutes an exclusion of the very possibility that organization might itself have causal power. If organization is "just" stage setting for "genuine" causal interactions, then emergent causality, which is supposed to emerge precisely in higher level organization, is ruled out by assumption.

That assumption, in turn, is supported by the basic metaphysical assumption that the world is constituted out of particles. If particles are the basic metaphysical constituents, then causality resides in those constituents, and all causality is the working out of the causal interactions among those constituents. Organization is not even a candidate locus for causal power. The conceptual possibility that organization might yield new causal power, the possibility that the notion of emergence is based upon, is delegitimated by the particle assumptions made in the underlying framework metaphysics. It is certainly clear that if particles are the only things that can possess causal power, then organizations of particles cannot possess causal power beyond that of the (interactions of the) particles themselves. The particle framework, in excluding organization even from consideration as a potential locus of causal power, begs the question against emergence.

But the metaphysical assumption here is that a particle metaphysics is adequate, or at least legitimate, as a framework for our world. And that is false. There is more than one line of considerations that yield this conclusion. First, consider that a pure point particle framework yields a world in which nothing ever happens: point particles have zero probability of ever hitting each other.[7]

We might consider particles that interact via some sorts of fields—and this is close to a common naïve conception today—but such a framework must acknowledge fields as having genuine causal power, and fields have whatever causal power they do in part in virtue of their organization. The delegitimation of relations as bearers of causal power, therefore, is undone.

Still further, the best contemporary physics demonstrates that there are no particles at all. The fundamental constituents of the world, according to quantum field theory, are dynamic quantum fields in a dynamic space-time. Quantum fields manifest particle-like properties in virtue of their interactions being constrained to occur in multiples of fixed quanta, and the conservations of those quantized properties

Footnote 6 continued
of emergence that are constituted rather than just supported by organization (e.g., "they supervene on specific mereological *configurations* involving these micro-properties—for a rather obvious and uninteresting reason: they *are* identical with these micro-configurations", Kim 1998, pp. 117–118; Kim 2005; see Campbell and Bickhard in preparation).

Such changes in definition, however, cannot do any metaphysical work in themselves, so Kim's earlier argument against causally efficacious emergence still stands. It is, in my judgement, one of the most powerful and revealing anti-emergence arguments available.

[7] For consideration of the conceptual possibility of non-point particles, see Bickhard (2000c, in preparation). Put simply, non-point particles would have to have boundaries, collision dynamics, internal structure, and internal dynamics, and those cannot be completed in ways that are coherent and consistent with relativity.

(Aitchison and Hey 1989; Ryder 1985; Weinberg 1995). The quantization is reminiscent of particles, but it is in fact a quantization of wave-like processes, not particles. This is akin, and mathematically related, to the quantization of the number of wavelengths in a guitar string (e.g., Zee 2003). But there are no guitar sound particles, and, similarly, there are no physical particles either (Aitchison 1985; Bickhard 2003c; Brown and Harré 1988; Cao 1999; Clifton 1996; Davies 1984; Halvorson and Clifton 2002; Huggett 2000; Kuhlmann et al. 2002; Sciama 1991; Weinberg 1977). The heritage of Democritus and Empedocles has run its course.

## 1.3 Process

And the manner in which it has expired points toward its replacement, both in physics and in metaphysics: *process*—Heraclitus, if you are so inclined. Quantum fields are inherently processes,[8] and general relativistic space-time is also inherently dynamic. It is, of course, possible, even likely, that contemporary physics will itself be overturned in the future, but it is clear that there cannot be any return to a particle framework: there are multiple phenomena with strong empirical support that are inconsistent with the locality and individuation of particles.

Process, in fact, is now the dominant language of science. Every science has progressed beyond an initial conception of its phenomena in substance terms to understanding that they are in fact process phenomena. Fire is no longer modeled in terms of the substance phlogiston, but instead in terms of the process of combustion; heat no longer in terms of caloric, but in terms of random kinetic processes; life no longer in terms of vital fluids, but in terms of special kinds of far from thermodynamic equilibrium processes. And so on. Every science, that is, with the exception of the sciences and philosophies of mind and persons. Here substance and structural views are still dominant.

The shift to a process metaphysics, however, induces major changes in our overall framework of assumptions:

- First, *change* becomes the explanatory default, and it is stability that requires explanation. Similarly, processes, unlike atoms or the "stuff" of substances, do not have inherent boundaries, and boundaries too, therefore, must be explained, not assumed.
- Second, processes have their causal powers in virtue of their organization. Organization cannot be delegitimated as a possible locus of causal power without eliminating all causality from the universe. But, if organization is a potential locus of causal power, then so is higher level organization. In particular, there is no metaphysical block to the possibility of emergent causal power in new organization.
- And third, if emergence is a metaphysical possibility (Bickhard 2000c, 2004a, forthcoming-b, in preparation), then the door is open to the possibility that

[8] There is some discussion about this, and a few alternative proposals are on offer (e.g., Clifton 1996; Kuhlmann et al. 2002; Teller 1996), but the mathematics is straightforwardly a mathematics of oscillatory processes (Weinberg 1995; Zee 2003), and a non-particle process perspective is the overwhelming consensus (the list of references in the previous paragraph could be multiplied many-fold, but see especially Halvorson and Clifton 2002; Huggett and Weingard 1996; Huggett 2000).

normativity and mind are emergent.[9] This would undo the two realm metaphysical split that has persisted for over two millennia.[10]

## 2 Emergence and normativity

A process metaphysics forces explicit consideration of stabilities of (patterns of) process—persistence through time cannot be assumed. There are at least two very broad classes of stabilities that I will consider, one of which is the foundation for the emergence of normativity in the form of normative function.

One kind of stability of process organization is that in which an instance of the organization remains stable unless some above-threshold amount of energy impinges on it. Such *energy-well* forms of stability constitute major portions of our world—a canonical example is an atom. Such stabilities can persist for cosmological durations. An important property of energy-well stabilities is that they can be isolated from their environments without disturbing that stability: they will happily go to thermodynamic equilibrium and remain in their "energy-well".

This property is in strong contrast to a second kind of stability of process organization: stabilities of processes that are far from thermodynamic equilibrium. Far from equilibrium processes cannot be isolated, else they go to equilibrium and cease to exist. The stability of a far from equilibrium form of process is dependent on its being maintained in its far from equilibrium condition. Such maintenance may often be accomplished from sources external to the process, such as pumps supplying a continuous flow of chemicals into a vat, perhaps for the sake of investigating the kinds of self-organization exhibited.

Some kinds of processes, however, self-organize into forms that make contributions to their own stability. In that sense, they are *self-maintenant*. A candle flame serves as a canonical example. A candle flame maintains above combustion threshold temperature, induces convection, which brings in fresh oxygen and gets rid of waste, vaporizes wax in the wick for combustion, and melts wax in the candle so that it can percolate up the wick. A candle flame is self-maintenant in several senses.[11]

---

[9] A process based model of the emergence of mental phenomena would clearly not be a substance dualism, nor a substance monism. Some readers might be tempted to interpret it as a kind of property dualism, but, insofar as emergent mental properties could be *explained* in terms of process organization, such a model would not constitute either a C. D. Broad style emergence model (McLaughlin 1992) nor a mysterian property dualism. Whether such a model would count as a monism of *any* kind would depend on whether dynamic space-time and dynamic quantum fields are to be counted as one kind (process) or several kinds (of process). Or whether or not emergent properties are to be considered of different kind than the process organizations out of which they emerge—if so, then there would be very large multiplicity of kinds: our entire familiar world, and much that isn't so familiar, is emergent in organizations of process. Most basically, in a process framework, this group of issues loses much of its philosophical motivation: they are posed within a Cartesian substance framework, and largely evaporate when an attempt is made to transplant them into a process framework. When stability, unity, and boundaries become themselves emergents of certain forms of process organization, classical contrasts no longer apply.

[10] It is this third issue that likely explains why studies of the mind are late in making the shift to process: mind (and biological function) are the realms in which issues of normativity and intentionality are central. Shifts to process in other sciences have not had to address these issues.

[11] Note: A candle flame is far from equilibrium relative to its environment—this is a relational property, and, therefore, does not have any "individual" level supervenience base (Teller 1992). A candle flame has no

A candle flame, however, can only do one thing—burn. It has no options and cannot select among options. If it runs out of wax, for example, there are no alternatives that it has the capacity to select, that might correct this threat to its continued existence. More sophisticated self-maintenant systems, however, do have options and can make selections among them in accordance with changing conditions in their environments in order to correct or compensate for those changing conditions.

Such systems maintain the property of being self-maintenant through changes of conditions that may require different activities for self-maintenance in those differing conditions. In that sense, they are *recursively self-maintenant*—they maintain self-maintenance. Here a canonical example is Don Campbell's bacterium that can swim so long as it is headed up a sugar gradient, but tumble if it finds itself oriented down a sugar gradient (Campbell 1974, 1990). Swimming up a sugar gradient contributes to its self-maintenance, but swimming down a sugar gradient does not. The bacterium can differentiate, detect, the two conditions and select swimming or tumbling accordingly.

Self-maintenant and recursively self-maintenant systems constitute the low end, the simple end, of a graded hierarchy of forms of *autonomy*. They are autonomous in the sense of being able to recruit and manipulate (themselves in) their environments so as to (contribute to) maintain(ing) their own existence (Christensen and Bickhard 2002). In complex cases, this can be a recursive self-maintenance across vast ranges of potential environment—it is an unbounded (and partially ordered) hierarchy. All living systems are autonomous.

The difference between energy-well stabilities and autonomous stabilities turns on a fundamental asymmetry in their underlying thermodynamics. Energy-well systems do well when isolated, but autonomous systems cannot be isolated for long because they require maintenance of their far from equilibrium conditions. This thermodynamic asymmetry underlies an asymmetric emergence, and, in fact, an emergence of normativity in the form of normative function. Reversal of all three of the substance-metaphysics consequences mentioned above is involved here.

### 2.1 Normative function

Far from equilibrium processes require maintenance in order to be stable, and such maintenance is functional relative to the stability of that system—it *serves a function* insofar as it contributes to that stability. This is the core notion of function.

Such contributions are functional relative to the process stability that they serve, to the maintenance of the existence that they help support. They are normative for that

Footnote 11 continued
fixed mereological base—the particular atoms are always and necessarily changing. If the candle is moved, then the incoming flow of oxygenated air will also change, but the candle flame will remain as a temporally continuous particular organization of processes—so the flame is not fixed to a particular region. There is no clear boundary for the flame. And so on. In multiple ways, such process organizations create difficulties for many versions of supervenience (Kim 1993; McLaughlin and Bennett 2005).

stability. They help create and maintain a dynamic steady state, and are useful relative to that maintenance and creation.[12]

### *2.1.1 Etiological models of function*

This model of function differs in several fundamental ways from the dominant etiological approach, and I will develop it in further detail by reviewing some of those differences. So, I turn first to outlining some of those differences.

The central intuition of the etiological approach to function is that an organ in an organism has some particular proper function insofar as it has the right kind of evolutionary history—in particular, that there has been a sufficient evolutionary past of selections of that organ for doing what is now its function. A kidney, for example, has the function of filtering blood because the ancestral kidneys to this one were selected for doing just that.

This is a design intuition. Kidneys have been "designed" for filtering blood by evolutionary selection, and it is in virtue of that "design", of that selection history, that kidneys have that function. Evolutionary "design", obviously, is not deliberate nor is it supposed to be in this approach. Instead, it is the selections occurring in evolutionary history that are equated to design. Kidneys have the function of filtering blood in virtue of the fact that they have been "designed" to do that, have been selected for doing that.

There is a strong appeal to this explanation. Clearly, kidneys exist because they have in general accomplished the filtering of blood, and, in that sense, filtering blood is what they are "for"—it is their function. Furthermore, this perspective makes sense of the notion that a kidney is "supposed" to filter blood even if this particular kidney is not doing that—this particular kidney is *dys*functional. If this model works, it has solved the problem of the emergence of normativity (normative function) and has somehow avoided Hume's argument against deriving norms from facts.

### *2.1.2 Epiphenomenality*

It does have, however, some strongly counter-intuitive consequences. Consider, for example, a thought experiment in which a lion springs into existence in the corner that, by assumption, is molecule by molecule identical to a lion in the zoo (Millikan 1984, 1993). The lion in the zoo has organs with appropriate evolutionary histories, and, therefore, that have functions. The science-fiction lion, however, has organs with no evolutionary histories at all, and, therefore, not the right kinds of evolutionary histories, and, therefore, absolutely no functions.

A non-science fiction version of this problem occurs during evolutionary histories in which a new organ accomplishes something useful for the organism or an "old" organ accomplishes something new that is useful. Such occurrences must happen in the beginning of the evolutionary selection histories that eventually generate particular organs with particular functions: until the selection history is sufficiently extensive,

[12] The point here is not that stability is itself normative. Rather, it is that stability (here) is the property in relation to which normative function emerges—contributions to stability contribute to that stability. And such contributions make a difference in how the world proceeds.

there is no function, but the usefulness must be there in order for there to be anything to be selected in the first place.

The focus of my criticism here is that we have cases, both science fictional thought experiment and real, in which one organism has the same dynamical properties as another, but one has functions and the other doesn't. The lion in the corner is dynamically identical to the lion in the zoo, but one has functions and the other doesn't. The early "usefulness" is as real a dynamic consequence as the later "function" but one is deemed to be functional and the other not.

Function, in this view, is dynamically—causally—epiphenomenal. It makes no difference to the causal or dynamic properties of an organism whether or not its organs have functions. Etiological models thus fail to naturalize function. Etiological history explains the etiology of something, but it does not constitute any of the causal or dynamic properties of that something. Etiology cannot constitute the dynamics of what it is the etiology of (Bickhard 1993, 2004a, forthcoming-a, forthcoming-b; Christensen and Bickhard 2002).

#### *2.1.3 Circularity*

Furthermore, the "usefulness" that must be selected for early in the selection history of a "function", according to etiological models, is itself already normative—if it were not, there would not be anything to select for. The etiological approach considered just as a model of the emergence of normativity is, therefore, circular: normative function is derived from normative usefulness, but normative usefulness is, in this model, not supposed to be normative. It is the "design" by selection that is supposed to generate normativity, but selection is already normative (Christensen and Bickhard 2002).

This "usefulness" is, as I have already contended, simply serving a function relative to the stability of an autonomous system. Etiological models, then, not only leave function epiphenomenal, they also presuppose normativity in the sense of contributing to the continued existence of a far from equilibrium system. They presuppose autonomy-based—recursive self-maintenance based—notions of serving a function.

Note that serving a function in this sense is a property that is constituted in the current dynamics of the system. It is not constituted in its history, however much it may be the case that the existence of this system cannot be understood without understanding its history. Serving a function in the sense of participating in and contributing to autonomy, therefore, is not epiphenomenal.[13]

### 2.2 Serving a function versus having a function

Another difference between etiological approaches to function and the autonomy model concerns what is taken to be the primary focus of explication. Etiological models focus on what it is to *have a function*, while the autonomy model focuses on what it is to *serve a function*. In either case, the second element of the pair requires further

[13] There are other challenges to the autonomy model of function. One interesting challenge has to do with the place of accidental contributions to stability. I address this issue (and others) in Bickhard (in preparation).

explication, but it makes a significant difference which is addressed first, primarily because that sets the bounds of the overarching framework within which to develop further explications. At issue is which is the conceptual genus within which the other can be explicated as a (derived) species.

In the etiological case, serving a function happens just when something that has a function succeeds in accomplishing the consequences that satisfy its function. A kidney serves the function of filtering blood when it in fact filters blood, and does so because it has the function of filtering blood. Note that, in this framework, no function can be served unless something *has* that function.

It requires a little more work to understand having a function in terms of serving a function—though what is required turns out to be of central importance for modeling other phenomena as well—but one advantage of taking "serve a function" as the focus of explication is immediate: functions can be served even if nothing has those functions. This opens functionality to a much broader realm of phenomena, and is, therefore, a more natural framework for modeling function.

#### *2.2.1 Functional presuppositions: the autonomy model of having a function*

The existence and maintenance of an autonomous system requires that particular functions be served, that particular self-maintenant contributions be accomplished. The specifics of such requirements will vary depending on the nature of the autonomous processes, but the far from equilibrium nature of the processes requires that that nature be maintained, and, therefore, that functions be served.

The successful serving of a function for some organism by a sub-process or part of that organism may depend on other conditions holding. Such other conditions can be of various kinds and take various forms. They may, for example, be environmental and ambient, such as that the atmosphere contain sufficient oxygen. Of particular importance here is that they might be conditions that are themselves created or maintained by other sub-processes or parts of the organism. The creation or maintenance of those conditions, therefore, is itself serving a function—the function of supporting the serving of further functions by subsequent processes and parts.

It is the converse of this relationship that I wish to focus on here: the serving of one function by some process or part may require, and, in that sense, presuppose, that other conditions hold. Insofar as those other conditions are created or maintained by some other process or part of the organism, the *first* process or part *functionally presupposes* that the *second* is successfully serving *its* function. It is a functional presupposition in that the first function cannot be served by that process or part (or at least not served as well) unless that second functional consequence is in fact created or maintained. The serving a function in the first case presupposes that the second function is being served.

Insofar as the presupposed serving of a function involves a presupposed location or part of the organism, then there is a functional presupposition that that part will serve its supporting function for the first serving of a function. In that sense, the part *has a function* by the presupposition(s) of other functional parts and processes. The functions of distributing food and oxygen cannot be served unless a carrier of appropriate constitution—e.g., blood—is located in a circulatory system and is in fact circulating in that system. The circulatory system, thus, *has* the (a) presupposed function of supporting blood circulation. The function of circulating blood, in turn, cannot be served unless something at the location of the heart serves the function of pumping—the heart, therefore, *has* the circulation of the blood as its presupposed function. And so on, in what is generally a very complex organization of interdependencies, presuppositions.

Having a function, therefore, is constituted in being presupposed to serve that function by the rest of the autonomous system. For now, the conclusion is that having a function *can* be modeled in terms of serving a function when the relationships of functional dependency and presupposition among various functions to be served are taken into account.

Functional presuppositions, it turns out, also have fundamentally important additional deployments.

## 3 Representation

In particular, special kinds of functional or dynamic presuppositions constitute the foundations for the emergence of representation. I begin the discussion of representation by addressing several alternative models of representation, arguing that they do not succeed in naturalizing representation, and then turn to the interactive model of representation.

### 3.1 Critical tools

Each of these models has its own particular problems, as well as multiple issues that are widely shared among them, and I will mention some issues in each of these categories, but the primary focus will be on how each of these models fails to meet two core criteria for a successful naturalistic model of representation. I turn now, therefore, to developing some of these "tools of criticism" for the representation discussion.

#### *3.1.1 General encodingism critiques*

I begin with some of the more general critiques. Many representational models hold that representation is constituted in some special (encoding) relationship between the representation and the represented. Typically, this special relation is thought to be causal, nomological, or informational.

*What are encodings?* I argue that representation cannot be *foundationally* constituted as encodings, and call the presupposition that it is *encodingism*. The problem is not that encodings do not exist—they clearly do, and are useful, even essential, for some purposes—but, rather, that encodings are a necessarily derivative kind of representation, and must ultimately be derivative from some other form of representation. In particular, encodings cannot provide representations of anything for which there is not already some representation available; for example, they cannot furnish (non-derivative) representations of the environment in mental processes. Encodings cannot cross the boundaries of epistemic agents because those agents must already have representations of their worlds in order for encodings to be derived.

Encodings are stand-ins, stand-ins for other representations. Consider, for example, Morse code, in which "…" stands-in for "s". Morse code is useful because dots and dashes can be sent over telegraph wires while characters cannot. Similarly, binary codes can be manipulated in computers. But the crucial point for current purposes is that neither Morse code nor computer codes can exist as such except insofar as various epistemic agents already know about, already represent, the dot and dash patterns, the characters, and the stand-in relationships between them, and similarly for computer codes. In other words, encodings change the *form* of representations, but borrow the *content* from elsewhere, which entails that, in order for encodings to have content, that content must already be available elsewhere.

Morse code and computer codes are conventional, but nothing about the necessary derivativeness of encodings turns on that. Consider a neutrino count in a mine that encodes properties of fusion processes in the sun. This is a strictly natural (causal, nomological, informational, etc.) relationship, but it is still the case that the neutrino count encodes those fusion processes only insofar as various physicists already know about, already represent, the neutrino counting, the fusion properties, and the relationships between them. The neutrino counts per se do not announce that they are in any kind of encoding relationship with anything, and certainly not what any such relationship might be with. Similarly, optic nerve activities per se do not announce that they are in any kind of encoding relationship with anything, and certainly not what any such relationship might be with.

Encodings are stand-ins, then, and as such cannot be the only form of representation. There must be some other form of representation for which encodings can be such stand-ins. The assumption that all representation has the form of encoding—encodingism—is false.

*Too many correspondences.* One immediate problem for encodingism is that, for any instance of a(n informational, nomological, etc.) correspondence that is presumed to constitute an encoding correspondence, there will in general be innumerable associated instances, and it is at best seriously problematic, and at worst circular, to model how some special one among these instances is the representational instance. Consider, for example, a supposedly representational visual causal relationship between some brain activity and a table. Any instance of this correspondence will be accompanied by further causal relationships between the brain activity and activities in the retina, the light patterns in front of the eyes, quantum activities in the surface of the table, the table a moment ago (in fact, the light relationship is with the table a "moment" ago), the table yesterday, the manufacture of the table, the generation of the materials that the table is made from (generally, in some past supernova), activities in the sun in which the light that supported the growth of the plants that formed the materials for the table was generated (if it is made of wood or oil), and so on all the way back to the Big Bang. These are all causal relations; which one is representational?[14]

[14] There have been attempts to address this issue. I will not address them specifically, except to point out that they are all framed within the perspective of an *observer* of the organism and its environment, not from the perspective of the epistemic agent itself (e.g., Smith 1987; Fodor 1990a,b, 1991). They thus do not address either representations for the epistemic agent per se, nor for the observer per se, but, at

*The copy argument.* Another issue is outlined in Piaget's copy argument: our representations cannot be in any sense copies of our world, because we would have to already know the world in order to construct our copies of it (Piaget 1970). This is one aspect of a common problem: attempts to account for representation generally either fail outright to capture representational normativity, perhaps even fail to address the issue—or are at root circular: they in some way presuppose the normativity that they attempt to account for.[15]

*Emergence.* Such circularities are a manifestation of another problem: they fail to account for the emergence of representation. Representation did not exist some thirteen billion years ago, and it does now. Therefore, it has to have emerged. Any model that cannot account for that emergence is at best incomplete, and any model that makes such emergence impossible is thereby refuted.

*Innatism and emergence.* This issue can be elaborated in ways that can make the underlying issue obscure. Fodor's argument for the innateness of representation (Fodor 1981), for example, turns on noting that models of learning do not account for the generation of new (emergent) representations, but instead constitute models of confirmation (or disconfirmation) of representations that are "molecularly" combined from previously available representations. He may be right concerning the available models of learning, but the conclusion that the sufficient base set of atomic representations must be innate does not follow: if representation emerged in evolution, then there must be some dynamic process by which such emergence is possible, and Fodor provides no argument that the brain cannot engage in the right dynamics to generate emergent representations—either in learning or development, or perhaps even in real interactive time. In fact, what the argument shows more strongly is:

> I am inclined to think that the argument has to be wrong, that *a nativism pushed to that point becomes unsupportable, that something important must have been left aside*. What I think it shows is really not so much an a priori argument for nativism as that *there must be some notion of learning that is so incredibly different from the one we have imagined* that we don't even know what it would be like as things now stand. (Fodor in Piattelli-Palmarini 1980, p. 269).

That is, there must be some model of learning (development) that can account for emergent representation, not just for the confirmation of previously available representation.

Footnote 14 continued
best, only for the agent-as-construed-by-the-observer (Bickhard and Terveen 1995; Bickhard 1993, 2004a, forthcoming-a, forthcoming-b). I will have more to say about this below.

[15] Note that "accounting" for the normativity of an agent's representation in terms of the (normative) representations, explanations, analyses, and so on of an observer is a common kind of such circularity.

*3.1.2 Two criteria: system detectable error and internal relations*

It is possible to undermine most models of representation with a small fixed set of critical tools because they still manifest their debt extending back to the substance metaphysics of the Greeks. Normativity is the focal problem, and the metaphysical split into two realms generates labyrinthine mazes of subproblems in trying to overcome it. Ultimately, so I argue, the split must be overcome via models of the emergence of normative phenomena out of non-normative phenomena, and, toward that end, the entire substance/particle framework must be reversed all the way back to a process metaphysics. A process framework does permit models of emergence, and, in particular, models of the emergence of normative function and representation.

*System detectable error.* One criterion that tends to focus these issues is that of the possibility of system detectable error—accounting for the possibility of an epistemic agent detecting, however fallibly and occasionally, error in its own representations. One indication of the depth of this criterion is that it is equivalent to overcoming or avoiding the radical skeptical argument, an argument that has withstood some considerable attempts to eliminate it (Popkin 2003; Popkin and Stroll 2002; Rescher 1980; Weintraub 1997). The radical argument is that no one can determine the truth or falsity of his or her own representations because to be able to do so would require that the individual step outside of him or herself to be able to compare what the representation represents about the environment with what is actually in that environment—no one can in fact step outside of him or herself, therefore no one can determine the truth value of his or her own representations. We cannot detect our own errors.

This radical skeptical problem is often dismissed or ignored with an attitude of "It does not seem to be solvable, therefore we need to move on and pursue other issues". But the problem is not just an "armchair" problem; it cannot be safely ignored because it has devastating consequences in accounting for related mental phenomena. In particular, if system detectable error is not possible, then error-guided behavior is not possible and learning is not possible (Bickhard and Terveen 1995; Bickhard 1993, 2004a, forthcoming-a, forthcoming-b, in preparation).

The possibility of system detectable error *must* be accounted for in any satisfactory model of representation. Representation must be normative—have truth value—for the epistemic system itself.

*Internally related content.* A second crucial criterion is that representational content must be internally related to representation. Elaborating this point requires first that the notions of internal and external relations be explained. The distinction is not one of a physical sense of internality or externality: an internal relation is internal in a sense of being necessary in order for one or more of the relata to be what it is (Hylton 1990; Bickhard 2003a, 2004a, forthcoming-b, in preparation). One example would be an arc of a circle, which cannot be that arc without bearing a specific relation to the point that is the center of the circle.

Notions of internal relations were prominent a century ago, but have become uncommon. The British idealists Green and Bradley made central use of internal relations, holding that everything is internally related to everything else in a kind of holism.

Russell reacted strongly against internal relations, and the holism that they had been used to generate, but could not avoid them totally—e.g., his type relations are internal relations (Hylton 1990).

One aspect of the idealist's framework that Russell was particularly exercised about was that they held that representations were symmetrically internally related to the represented (a simple consequence of internally relational *holism*). This had the consequence that a change in a representation entailed a change in the represented. Russell reacted against this rather counter-intuitive consequence with an externalism involving, as much as possible, only external relations. He maintained, however, the framework of a two-part model of representation—representation and represented—and rejected (at least for some time) anything like Frege's Sense or content (Hylton 1990).

The near demise of internal relations occurred with Quine's rejection of all things that did not fit with his austere physicalistic ontology (Gibson 2004; Glock 1996, 2003; Orenstein 2002). In particular, rejected "violations of austerity" included properties of the realm of mentality, such as intensions, and much of modality (except perhaps logical necessity), such as any notion similar to "essence". Internal relations are relations of necessity, and, in general, not just logical necessity, and, thus, they fell to this resurgence of substance- or particle-physicalism.

I argue, however, that some issues cannot be understood without recourse to internal relations. One of these is representation. In particular, if a representation is externally related to its content,[16] then it follows that there is nothing about that representation that requires that it have that content, or any content at all. Consequently, if there is to be externally related content, it must be *added to* the representational element, and it must be understood to have so been added by any epistemic agent taking that representation as a representation. But, if such externally related contents are all that are available, then there cannot be any account of content per se. Providing content from some other representation can certainly be done—real encodings are defined that way—but if the further representation also bears externally related content, then we encounter the familiar infinite regress of interpreters interpreting in terms of other encodings (or symbols) which also bear only externally related contents, which, therefore, must be provided by still some further representation, and so on unboundedly. Any such interpreter is an undischarged homunculus circularly providing the content that is supposed to be accounted for in the model.

In general, assuming externally related representational content renders content metaphysically unrelated to anything that might constitute a representation. Therefore approaches based on external relations cannot address any issues of emergence. They lead to an infinite regress of interpreters in attempting to find an ultimate ground of content that can be used to define all other representations.

On the other hand, if content is *internally* related to representation, then whatever constitutes such a representation cannot be what it is without bearing that content, and

[16] If the model is a two part externally-related model, with no distinct content, then the problems outlined above recur with respect to the representation-represented relation, but, if that relation is taken as an internal relation, then the consequence that Russell reacted against resurfaces. Representation cannot be just a two part relation (Bickhard 2003a).

no content has to be added or interpreted in any external manner. The infinite regress of interpreters never gets started.

Taking a historical perspective for a moment, note that a central theme amongst the Greeks was the intuition that “like represents like”. Aristotle’s version of form-representing-form—as in the metaphor of the signet ring leaving its form behind in wax—was arguably an instance of an internal relation: identity. It nevertheless fails, for example, to account for the possibility of error, and certainly not of system detectable error. We find contemporary descendents of such notions in models of representation as iso-(or homo-)morphisms of structure between representation and represented. In such models, the internality of the representational relationship is lost, and multiple additional problems appear.

More generally, any such internality was lost at least by the time of Locke in the “fracture of the forms” (Campbell 1992), in which form-representing-form was replaced with explicitly external relations as models of representation, such as causality. Sophisticated descendents of these notions are also a part of the contemporary scene: e.g., nomological or informational (covariation) relations as constituting representational relations.

Contemporary models of representation, then, are all examples of attempting to make use of external relations to account for representational relations. As such they not only encounter the problem of an unbounded regress of interpreters, but also run afoul of the system detectable error criterion. One requirement to be able to detect error in one’s own representations is that the content be accessible, but, with nothing but externally related content available in the models, there is no way in which an organism can access its own content in order to even begin to assess that content in the current environment.

## 3.2 Some contemporary literature

I turn now to a brief examination of some contemporary models of representation, specifically, those of Millikan, Dretske, Fodor, and Cummins.

### *3.2.1 Millikan*

There is an overall architectural similarity between Millikan’s model of representation and the interactivist model being outlined here: both initially develop a model of normative function, and then a model of representation based on that of function. Beyond this macro-architectural level, however, there are significant differences. As already indicated, the models of function differ; as we shall see, the models of representation differ, and the relationships between function and representation differ.

Millikan’s model of function is etiological—having a function is constituted in having the right kind of evolutionary history. In consequence, it is epiphenomenal. Representation is itself, in this model, a kind of function, and, therefore, it too is epiphenomenal.

Content—what a representation is supposed to represent—is similarly constituted in the history of a representation, and, therefore, it is not only epiphenomenal, but it is also externally related to any dynamics in the current organism, and correspondingly inaccessible to that organism.

System detectable error, however, requires that content be compared to what is being represented, which, in turn, requires that content be accessible to the organism. So an etiological model doesn't provide the first basis for such error detection: the content is beyond reach.[17]

Even if content were accessible, comparing that content to what is being represented in the environment would require epistemic access to that which is being represented, but that is the original problem of representation again. This is precisely the circularity revealed by the radical skeptical argument. It recurs for every one of the models under consideration. System detectable error is not possible: content is neither accessible, nor comparable.

### *3.2.2 Fodor*

Fodor's model is a version of an information semantics: representational relationships are constituted as some form of informational relationship (Fodor 1990a,b, 1991, 1998, 2003). Regarding the problem of representational error per se, Fodor is in a more difficult position than Millikan: for Millikan, content is constituted in past history, and, therefore, is independent of what is currently being represented, and there is no conceptual difficulty in modeling the possibility that that content fails to match what is currently represented.[18] For Fodor, however, representation is constituted in informational relationships with what is being *currently* represented, so it is not immediately clear how error per se is possible. If the crucial (informational in this case) relationship exists, then the representation exists, and it is correct, while if the crucial relationship does not exist, then the representation does not exist—existence or non-existence of the crucial relationship would seem to be the only two possibilities available for the model, but there is a third representational possibility that must be accounted for: the representation exists, but it is false.

Fodor has proposed an ingenious attempted solution to this problem: a kind of asymmetric dependency between true instances and false instances. The basic idea is that the false representations—e.g., a representation of a cow, but applied to what is in fact a horse on a dark night—are dependent on the true instances, in the sense that the cow representation would never apply to the horse on a dark night if it did not apply to, or be evoked by, a genuine cow—so there is a dependency of the false on the true—but the dependency is not reciprocated in that it is quite plausible that the cow representation is evoked by cows even if it would never be evoked by horses. False representation is parasitic on true representation.

There are, however, counterexamples to this property as an account of the possibilities of true and false representation. Consider a poison molecule that mimics a

[17] Note that this point is not simply that the content has its *origins* in the past—that would hold of most anything. It is a historical *metaphysics* that is at issue, not a historical origin. The two lions example as well as others demonstrate that nothing in the current organism can dynamically depend on that past qua past—including on content. Vastly different histories can yield identical current dynamics.

[18] Setting aside issues of access, comparability, epiphenomenality, and so on.

neurotransmitter. In both cases—the neurotransmitter and the poison—we have causal, nomological, and informational relationships between the responses inside the receiving neuron and the activities at its receptors. Furthermore, the fact that the poison molecule succeeds in docking in the receptor molecule is asymmetrically dependent on the genuine neurotransmitter being able to dock in that receptor, but that dependency is not reciprocated. This example satisfies the model, yet there is at best a functional error here, not an epistemic error. Asymmetric dependency perhaps captures a property of certain kinds of parasitic relationships, but it does not suffice to capture representation (Bickhard 1993, 2004a; Levine and Bichard 1999).

Furthermore, the content of a representation in this model is constituted in complex asymmetric dependency relationships between classes of counterfactuals concerning what would happen if multifarious possibilities were to occur. So, even if the model were to be able to account for error per se, it cannot account for system detectable error.[19] Content is external: it is what is actually on the other end of the nomological informational relationship (somehow packaged together with the asymmetric dependency conditions), thus not accessible and subject to the interpretive regress problem.[20]

And, as before, even if content were accessible, comparing it to what is in fact being represented would require independent access to what is being represented, and that is the representational problem all over again. Radical skeptical inaccessibility and circularity yet again.

### *3.2.3 Dretske*

Dretske's model, like Millikan's, is an etiological model, though in this case the relevant etiology is a learning etiology rather than an evolutionary etiology (Dretske 1988). Correspondingly, it is also epiphenomenal, and with externally related and inaccessible content. And, as is by now familiar, even if content were accessible, system detectable error would require comparison with what is currently being represented, which is the radical skeptical circularity.

Dretske's model also illustrates another issue. In a crucial sentence, he states that:

> C is recruited as a cause of M *because* of what it indicates about F, the conditions on which the success of M depends. (Dretske, 1988, p. 101, emphasis in the original).

Here, C is an internal state, M is some behavior, F is the conditions for the success of the behavior M, and the indication relation is equivalent to "is correlated with". Indication—a kind of informational relation—is the core notion in Dretske's model of representation.

What I would like to point out here about this sentence is that the "because" cannot be a causal "because": The indication relation between C and F is not accessible to the organism, neither the existence of such a correlation, nor what any such correlation might be with. So neither the existence nor the environmental end of such a relation

[19] And, like all of these models, there is no attempt in the literature to address system detectable error.

[20] Causality is a classic example of an external relation.

can have any dynamic or causal consequences in the organism. Such a relationship is determinable only from the perspective of an external observer of the organism and its environment together, only for an external homunculus who can provide the connection between C and F.

This is not an immediate problem for Dretske because he models representation not as a kind of phenomenon in an organism per se, but as a matter of the best explanation of that organism's activity. Representation is *attributed* if the crucial indication relationships are part of the best explanation of behavior. This requires an "explainer", even if idealized in some way, and some phenomena do in fact have such an attributional ontology—e.g., money or marriage.

But the presumption that representation has such an attributional nature encounters the problem of regress: the attributions or explanations of the person or persons engaged in explaining the organism's activities are being crucially adverted to here, which presupposes the existence of the representations of any such external observers. The external observers are the epistemic homunculi from which the normative content for the organism's alleged representations are being derived. But these representations of the external observers are not themselves accounted for, and any attempt to account for them in like attributional manner simply initiates the familiar regress of interpretive homunculi. Their status of being (physically) external doesn't change the basic problem. Among other consequences, the possibility of system detectable error is doubly precluded: any content is not only external in the metaphysical sense, it is also external in the physical sense—in the external observer's mind.

More generally, modeling representation as having any form of external attributional ontology encounters similar circularities and regresses, and does not succeed in dissolving or transcending the external relation or system detectable error issues.

### *3.2.4 Cummins*

Cummins (1996) proposes an interesting model of representation that is a relative of the "like represents like" tradition. In particular, structure (for the organism) is supposed to represent iso- or homo-morphic structure (in the environment). This framework prima facie shows promise in solving at least the problem of accounting for the possibility of error: if the "target" of a representation can be successfully identified independently of the representation(al content) itself, then the structure that constitutes the representational content might well fit or fail to fit the structure of the target. Content and "target" are pulled apart[21] and, therefore, accounting for the possibility of their differing is not per se problematic.

But this model too does not succeed. I will limit the discussion to two crucial ambiguities in the model, both of which can be illustrated with one of Cummins' examples. Consider a toy car into which a card can be inserted. The card has a groove in it that changes position on the card from front end to back end, and the car has a peg that fits into the groove and can follow it as the card is moved into the car. The shifts

[21] Both etiological models also share this property: if content is constituted in the past and the "target" of representation is in the present, then, at least to a first approximation, there is no difficulty accounting for the possibility that the content fails to fit the represented.

of the peg, in turn, can be used to control the turning of the front wheels of the car, thus steering it. If the groove has a structure isomorphic to the structure of a maze, for example, then the car can successfully run the maze, and structure would seem to represent structure.

The first problem is that structure in this sense is not an intrinsic property. It is relational. Suppose, for example, that the car did not trace the physical oscillations of the groove, but, instead, read changes in direction or strength of magnetization along the groove (if there were any). A completely different "structure" would, in general, be manifest. Structure, then, is not intrinsic, but is a relational matter—related to manner of read out. This point is not obvious in Cummins' discussion because he tends to focus on one of two kinds of cases: mathematical structures, which are fixed by stipulation, and examples that at first seem obvious, such as the car with the slot card.

The second problem has to do with the structure that is supposed to belong to the target. If the goal of putting the car into the maze is to hit the wall at a certain point in the maze, then the original card will not represent the target at all. Conversely, *whatever* the car does, that could be taken (or not) as satisfying "the" goal, and so the "representing structure" would constitute a true (or false) content. What counts as a target is a normative matter, and this normativity is not addressed in the model.

In this model, the presumed structural content is externally related to the supposed bearer of that structure because it is a matter of read-out relation, rather than an intrinsic matter, just what that structure is. There are some interesting issues to consider if the structure bearer, such as the card, is taken to be a part of the system so that the relational issues are themselves part of the system, but, in any case, the normative issue of what constitutes the target precludes system detectable error. Identification of the target is a normative issue, not just a causal issue, and accounting for that normativity and how it relates to the environment is unaddressed.

### 3.3 Desiderata for models of representation

In the course this examination, several desiderata for models of representation have come to light. One central criterion is that a model of representation must be able to account for system detectable error: again, if not, then it cannot account for error guided behavior and learning, and is thereby refuted. To be able to account for system detectable error, in turn, involves satisfying further criteria. Two are that the content must be functionally accessible to the organism, and that the content must be subject to comparison with whatever is being represented—which, therefore, must itself be accessible. In order for content to be accessible, it must either be provided from some other source of representational content (creating an externally related content) or it must be internally related to whatever constitutes representation. Since any other source of content must ultimately halt an infinite regress of provisions of content, there must be some ground of representing with internally related content. Representation with internally related content, in turn, entails that the creation of any such representation constitutes the creation of its internally related content, thus constituting emergent representation.

So, we have desiderata of:

1. system detectable error,
2. functionally accessible content,
3. functionally comparable represented,
4. internally related content, and
5. emergent representation.

*None* of these desiderata are satisfied by any of the standard models in the literature.[22] I will now turn to outlining the interactive model of representation and showing that it does satisfy these criteria.

## 3.4 Interactive representation

Processes that serve functions can involve functional presuppositions—conditions that support their serving the general function of self-maintenance. As discussed, when inherent in the functional organization of a system and concerning the supports from particular parts or organs, this constitutes a presupposition that those parts or organs will in fact serve that supportive function, and, in that sense, they *have* the function of *serving* that function.

### *3.4.1 Minimal representing*

But functional presuppositions are not confined to within the boundaries of an organism. In particular, a recursively self-maintenant autonomous system does things in and with its environment. In at least a minimal sense, it is an agent. It interacts. Its interactive activities are part of its autonomy; they constitute, in general, contributions to the (recursive) self-maintenance of the system. And those activities will, in familiar form, succeed in making such contributions under some circumstances and not under others.

When they are pragmatically successful in making such a contribution, they are fulfilling their function of contributing to autonomy. They are agentively true to their nature as activities helping to constitute that autonomy. When they do not, they are false, or unfaithful, to that nature (Campbell 1992).

Engaging in such interactions is functionally anticipatory that they will in fact succeed in contributing to the maintenance of the autonomous system. That is, the success or lack thereof, and the conditions upon which that might depend, are in the future of the interaction once initiated. As such, there is an implicit predication that "this" is one of those environments in which the initiated interaction will proceed as anticipated. That predication, therefore, might itself be true or false: the environment might or might not be among the supportive kinds.

Under some conditions, the activity will be functional, while under others it will not. Initiating the activity, therefore, presupposes that those supportive conditions hold.

[22] It might be argued that models of representation as structural iso- or homo-morphism satisfy 2 and 4, but, since it is not clear what the relevant structure is in any physical case, it is not clear how it is to be accessed nor what it could be internally related to.

This is a functional presuppositional relationship between the activities of the autonomous system and its environmental conditions, a presuppositional relationship that extends outside of the organism. The activity will only be functional, and, therefore, only be true to its nature, if those supportive conditions hold.

This, I propose, constitutes a minimal model of representing. Such anticipatory activities have truth-value. If they do not proceed as anticipated, they are false *and falsified for the organism*. That is, if the internal preparedness for engaging in the interaction is violated, if the actual course of the interaction proceeds beyond the bounds of what is functionally anticipated, then that anticipation is falsified—for the organism itself.

Much more of this model needs to be elaborated in order to show that it is a candidate for being able to account for all representational phenomena, but, before turning to some of those elaborations, let me examine what is already available with respect to the desiderata listed above.

- We have truth-value, relative to the autonomous nature of the system and its interactions.
- We have system-detectable error: if anticipations do not proceed as anticipated, they are falsified.
- We have accessibility: the functional preparedness of the system is functionally accessible physically internal to the system.
- We have comparability: what is anticipated is a range of possible futures for the functional processes of the system itself. So, they too are functionally accessible in a manner that is physically internal to the system.
- We have internally related content: no autonomous agent of this kind could engage in such interactive activities without involving the same functional presuppositions concerning supporting environmental conditions. The activity could not be what it is without being in that presuppositional relationship.
- And, finally, we have emergence of representational phenomena out of non-representational phenomena: the organization of an autonomous system that makes it a recursively self-maintenant system does not require any representational phenomena in order for it to exist or to come into existence. But once it does exist, the relevant nature of autonomy, the implicit predications concerning environments, and the presuppositions involved in those predications will all also exist. The emergence of representing and representational content can be accounted for.

### *3.4.2 More complex representing*

To this point, the model outlined is minimal—though it does satisfy the most difficult desiderata, criteria that are *not* satisfied by alternative models. In order to address some examples of more complex representing, I turn now to some of the resources available in the basic model.

The crucial differentiation, both conceptually and in phylogenetically, is between two aspects of the triggering relationship that we find in the bacterium—the triggering of swimming or tumbling depending on detected conditions.[23] More complex

[23] It is worth pointing out that detection per se neither constitutes nor requires representation.

organisms, such as a frog, may have multiple interaction possibilities available. There might be a fly in one location—i.e., a possibility of flicking its tongue and eating in that location—another fly in another location, and a worm in still a third. The frog must select among such possibilities, but, in order to be able to do so, it must have some sort of functional indications of those possibilities among which it can select. Both the selection process and the indication relationship are of central interest, but here I will focus on the indications.[24]

The first point is that such indications of interactive potentialities have truth-values in the same sense as do triggered interactions: they are anticipative and might fail. The second point is that, with such indication functions, the possibility of branching interaction potentialities emerges. The frog, in this example, has three potentialities available.

The third point is slightly more subtle. If the frog detects an opportunity for tongue flicking and eating via a factual detection of a fly, then there is a functional relationship in the frog system that *would* have indicated that tongue flicking and eating possibility *if* a fly *had* been detected—and that *conditional relationship* for creating such an indication is functionally present in the frog even when there is no fly in that location—even when the conditional is counterfactual. That is, there must be conditional preparednesses to set up indications of interaction possibilities, even when the conditions are not satisfied and so the indications are not activated.

At this point, we have the possibility of branching indications of interactive potentialities and of conditional indications of such potentialities. When such branched conditionals are able to iterate, so that one interaction may create the conditions for another (set of) interaction possibilities—I can get a drink from the refrigerator, but first I have to open the door—the organization of such conditionalized interaction possibilities can form a web. Perhaps a very complex web.

It is this web that constitutes the primary resource for the emergence of more complex representing. I call such a web the organism's *situation knowledge*—knowledge of the interactive potentialities of the current (extended) situation.

An organism's interactive potentialities do not remain constant. They change continuously in accordance with the organism's activities, as well as with processes and activities occurring in the environment. The situation knowledge web, therefore, requires ongoing updating and maintenance. For processes of situation knowledge updating and maintenance I have adopted the term *apperception*.

*Small objects.* More complex representing is of multifarious kinds, including representations of objects, various kinds of concepts, representations of processes, numbers, and so on. I will illustrate the resource of situation knowledge webs by outlining a model of small object representation.

Consider a child's toy block. It offers multiple possibilities of visual scan and manipulation. Any one of these interaction possibilities can be accessed from any other via appropriate intermediary interactions: e.g., a visual scan of the back of the block is reachable by turning the block over so that what was the back side is now visible.

[24] Processes of selection—more generally, processes and tendencies of interaction and development—open up the realm of motivation (Bickhard 2000b, 2003b, 2006, in preparation).

The organization of these interactive potentialities, therefore, will be an internally completely reachable sub-web of the overall situation knowledge web.

Furthermore, that internally reachable sub-web will itself have an important further property: it is invariant under a wide range of other processes and activities. The child can drop the block, chew on it, leave it on the floor and go into the next room, put it away in the toy box (it's at least a possibility!), and so on. The interactive potentialities that constitute the interactive nature of the block will remain invariant and reachable under all such changes, given appropriate intermediate activities, such as returning to and opening the toy box. It is not invariant under all changes, however. Crushing or burning the block, for example, eliminates the interactive potentialities.

This is basically Piaget's model of the representation of a toy block stated in interactivist terms (Piaget 1954). It is possible to borrow from Piaget's model in this way because both Piaget's model and the interactivist model are action- and interaction-based. They are both part of the general family of pragmatist models. I would not endorse all of Piaget's model, but it is a rich resource for interaction-based models—and thus radically unlike most models in psychology, cognitive science, and philosophy.[25]

*The number three.* One additional challenge to an interaction based model of representation that I will consider is that of abstractions, such as number. It might be granted that an interaction model could perhaps account for representations of external objects and processes, but where is the realm of interaction in which abstractions can be represented?

Consider a heuristic strategy servomechanism that might be constructed in a complex interactive system that would, when given a way of interacting with the environment and a goal, try that form of interaction three times before giving up and passing control on to some other process. Here is an instantiation of ordinal three. But a simple interactive system cannot interact with itself, at least not directly, so this (potential) property of an interactive system would seem to be beyond reach of interactive representation.

If, however, there were a second level interactive system that interacted with the first level system in much the same way as the first level interacts with the environment, then that second level could differentiate and represent organizational and processing properties of the first level such as "being an instantiation of three". That is, the interactive system itself can constitute a realm of instantiations of abstract properties.

The relationship of higher level interactive systems representing properties of lower level interactive systems iterates, and generates a hierarchy of potential levels of knowing, each level interacting with the next lower level, and potentially representing new properties emergent at that level—such as higher order invariances. I will not develop the model of how ascent through these levels can occur (see Campbell and Bickhard 1986; Bickhard 2006), but it is similar to, though in this case also strongly different from, Piaget's reflective abstraction (Piaget 2001). For now, the point is that there

25 And unlike the distortions of Piaget's work that have formed the basis for the anti-constructivist orientation of developmental psychology for the past three or four decades (Allen 2007; Chapman 1988).

is no aporia concerning the possibility of representing abstractions in the interactive model.[26]

The resources of the situation knowledge web and levels of knowing provide a powerful framework within which to approach complex and abstract representation. The two examples of a toy block and a number don't begin to exhaust all of the representational phenomena that ultimately must be addressed, but they do indicate that there are rich modeling resources to explore. The interactive model of representation, thus, accounts for basic normative and intentional properties of representation that are not captured by alternatives in the literature, and provides a framework within which more complex representational phenomena can be addressed. It is a viable candidate for capturing the basic nature of representation.

## 4 Some consequences for cognition

The interactive model of representation both constrains and enables models of further cognitive phenomena. Issues of cognition and representation are ubiquitous in psychological phenomena, so the consequences of an interactive approach are similarly widespread. Issues that have been addressed thus far in the development of the model include: perception, memory, motivation, concepts, learning, heuristic learning, emotions, consciousness and reflective consciousness, CNS processes and evolution, rationality, social ontology, language, persons, etc. Here I will outline some consequences for three phenomena: perception, learning, and language. The intent here is not to present anything like a complete model of any of these three, but rather to limn and illustrate some of the consequences of addressing such phenomena from within the interactive framework. Phenomena do not look quite the same from within this perspective.

### 4.1 Perception

Perception is standardly modeled as an input-processing phenomenon, with the inputs being some sort of sensory encodings generated or transduced by the sensory receptors (Carlson 2000; Fodor and Pylyshyn 1981; Bickhard and Richie 1983). But "transduction" is a causal process, and, therefore, there is no internal relation between anything occurring in the nervous system upstream of any such transduction and any kind of content that such nervous system processes might be supposed to carry. Consequently, there cannot be any emergent content generated along with the causal aspect of transduction processes—transductive "encoding" cannot generate the basic representations upon which further input processing are supposed to operate.

> It goes without saying—anyhow, it *ought to* go without saying—that encoding … [is] pie in the sky so far. [It is a] semantical notion, and—as things now

[26] This hierarchy of levels of potential interactive knowing has many additional important properties and consequences (e.g., Campbell and Bickhard 1986; Christopher and Bickhard 2007; Bickhard 2002).

stand—we haven't got a ghost of a Naturalistic theory about [it]. (Fodor 1987, pp. 80–81).[27]

Absent such transductive encoding emergence of internally related content, there would have to be some sort of homunculus to *add* content to the products of transduction, as well as a homunculus to *interpret* those encodings in terms of their externally related content. This initiates the classic regress of interpretive homunculi.[28]

But, if perception is not a matter of sensory encoding, then what is it? And what happens to the classic information flow from perception to cognition to action (including language)?

#### *4.1.1 Contact and content*

Representational content is basically future-oriented, anticipatory. But, if such anticipatory indications of interactive potentialities are to have any probability of being correct, they must be set up in ways that are appropriately sensitive to, modulated by, some sort of contact with the environment. Consider an interactive (sub-)system engaged in interaction with an environment. The internal course of the interaction will depend in part on the organization of the system, and in part on the environment being interacted with. The internal course and outcome of such an interaction, therefore, will serve to differentiate categories of environments from each other. For ease of discussion, consider such a sub-system that has only two possible internal outcomes, A and B. If the interaction ends in state A, then the organism is in an A-type environment—the type of environment that yields internal outcome A—and, if it ends in state B, then it is in a B-type of environment.

At this point, there is no content involved in being in internal state A or B. There is just a differentiation of two kinds of environments, with no characterization of those environments. So, a detection of an A-type environment is *just* differentiation, not a representation. Arriving at internal state A, however, can nevertheless potentially be useful to the organism. It might be learned, or hard-wired, for example, that, if state A is encountered, then an indication of the possibility of tongue flicking and eating of a particular sort can be set up. Such an indication *is* future oriented, anticipatory, and, therefore, involves content: it is about the current environment, and it could be true or false.[29] But, to reiterate, setting such an indication up should in general be contingent

[27] For a more recent quote, among many, making the same point: "Hume hasn't, in short, the slightest idea how 'the world' or 'the object' (or anything else) *could* cause an impression (and neither, of course, do we)." Fodor 2003, p. 121, footnote 10.

[28] For more thorough critiques of Fodor's attempts to capture the core of such a process, see e.g., Bickhard and Richie (1983); Bickhard (1993); Bickhard and Terveen (1995), and Bickhard (in preparation).

[29] The structure of possible outcomes of a differentiating interaction imposes a partition of the differentiation outcome classes on possible environments. Such partitionings replace standard conceptions of epistemology as based on correspondence—generating a *partition epistemology*—but not only are the partitions not necessarily unique to individual entities (or to entities at all), in themselves they do not represent at all, and do not have truth values. However, indications that, if a differentiation process has arrived at one partition category, then the system is also in some other interactive differentiation-category—indications of further potential interaction—do have truth value, and ground representation in general.

on having engaged in a prior differentiating interaction with the right kind of internal outcome.

Consider now the possibility of such a differentiating interaction that has no outputs (null outputs). This would be a strict input processing differentiation, but, insofar as such processes actually exist, it could still potentially serve to usefully differentiate environments in accordance with the internal outcome states that such input processing produced. It is assumptive models of such passive input processing that constitute the core of classical models of perception. Crucially, however, such classic models interpret the differentiating detections that are thereby produced as constituting representations of that which they have differentiated. They are taken to encode what they have differentiated. It is this step that yields the labyrinth of encodingism problems.

In the interactive model, in contrast, such differentiations are a necessary form of *contact* with the environment, and can serve as the functional basis for setting up anticipatory *content* about that environment, but the differentiations do not themselves constitute representations. Standard models, in other words, conflate contact and content.

More broadly, perception rarely involves such passive input processing, even as a first step. Perceptual interactions are those which are engaged in primarily for the purpose of making differentiations about the environment, which can then modulate the processes of apperception concerning the interactive potentialities in that environment. That is, perceptual interactions are those that are engaged in primarily to serve the function of apperception.

Stated this broadly, perception includes such processes as reading X-rays or sonar, or engaging in qualitative analysis in chemistry. More narrowly, there are major classes of interactive modalities that are primarily engaged in for apperceptive purposes for which we have physiologically differentiated and specialized sub-systems. These form the classic sensory systems: vision, hearing, touch, and so on.

It must be noted, however, that the distinction made here is *not* co-extensional with perception in the classic input processing model. First of all, "perceptual" interactions can involve any set of muscle groups: motion to pick up parallax, shifting one's head in order to see better what might be behind some leaves, moving the branches with your hand in order to see better what might be behind the leaves, and so on. Conversely, all interactions differentiate some things and change others, and even classically perceptual interactions can be used to change things in the world, not just to detect them. A shift of the eyes, for example, might be used to signal someone.

The information flow model of perception, cognition, and language is wrong from top to bottom. Information relationships, in the sense of a relation of correlation or covariation, play important roles, but they are control theoretic roles of being constructed in differentiations and being constituted in future oriented indications of interactive potentialities (Bickhard 1993, 2000a). Information is *not* in itself a semantic or representational notion.

Perception, then, looks different from an interactive perspective. It is not an input phase of information processing. It is, instead, a kind of interaction for primarily apperceptive functions, that is further differentiated in terms of several physiological differentiations of modalities of interaction that are mostly engaged in for their apperceptive support. The fundamental interface with the environment is interaction, with some kinds of interactions being more perceptual—more exclusively and more physiologically specialized for apperception—than others.

### *4.1.2 What is perceived?*

In this model, perceptual interactions are engaged in, but they are not themselves contentful. What is contentful are the processes of apperception and the processes of interaction guidance and selection on the basis of the apperceived situation knowledge. These processes constitute a flow of contentful activity that is inherently situated, from an embodied point of view, and, therefore, context dependent on that embodied situatedness.

To illustrate some of how this can work, I will address three examples of perceiving something and show how they can be understood from within this framework. The three are a child's toy block, a straight line, and red.

*Toy block.* The representing of a toy block has already been outlined: an internally completely reachable sub-web of the situation knowledge web that is invariant under various kinds of motions and locomotions. What I want to point out here is that such a representation of a toy block is primarily anticipative, even when someone is engaged in interaction with the block. The ongoing guidance of interacting with the block is in terms of that anticipative organization, and no actual interaction will exhaust all of the interactive potentialities.

This point holds even when the interaction with the block is very limited and primarily perceptual. If the block is seen with a visual glance, then (apperception of) the situation knowledge of there being such an organization of interactive potentialities is what constitutes the seeing of a block in that location. In particular, it is not the visual interaction per se that constitutes the experience of seeing the block; it is the apperceptive construction in the situation knowledge and any taking of that "block" into account in the guidance and selection of further interacting that constitutes that seeing. This process is contentful, situated, and so on.

Furthermore, *mis*-perceiving something as a toy block is similarly anticipative. It is to apperceive the potentialities of particular patterns and properties of potential interactions. Engaging in those interactions, perhaps even just a closer visual scan, might falsify those anticipations, thus falsifying the situation knowledge representation that there is a toy block there.

It is the point that contentful process—i.e., experience—is inherently anticipative that persists through the examples. This is a general point about experience, about the locus in which contentful experience is emergent.[30]

*Straight line.* Consider next perceiving a straight line—or, better, seeing a line as straight. This could, for example, be a boundary between light and dark, or between differing colors.

[30] This point together with the internal relatedness of such content in experience, and with the knowing levels, can dissolve a number of perplexities about consciousness (Bickhard 2005).

The projection of a straight line on the retina will not, in general, be itself straight. A boundary may cut some receptors in the middle, others just on an edge, and still others may be just missed by the boundary. But, if the line is scanned along its length, that somewhat irregular pattern of receptors will remain invariant. If the line is slightly curved, however, then some that were cut just on an edge may no longer have the boundary projected on them at all, and others that were just outside of the edge will now be hit by it. That is, if the line is not straight, then scanning along the line will *not* yield an invariance of pattern.

Note again that seeing a line *as straight* will involve *anticipating* that scanning along its length will yield an invariant pattern, and this will be the case even if the line is not scanned at all. What is seen—a straight line—is anticipative. It is a property of the apperceptive and situation knowledge processes.

*Seeing red.* This general point can be extended to phenomena such as seeing red. The distribution of primarily red sensitive receptors is not regular across the retina. It is densest in the central macula, decreasing toward the edge of the macula, and decreasing still more outside of it. So, looking at a red spot will evoke a particular pattern of "red receptions", depending on exactly where the red is projected on the retina. A shift of visual direction will change both the particular red receptors activated, and the pattern of them, again depending on where the new projection is located on the retina.

Crucially, variations in visual scans will change the patterning of receptions evoked by the red spot in ways that themselves involve an invariance. The changes in receptor patterns will form a mathematical group, complete with identity (don't change visual direction), composition (two shifts will always give the same result as some third shift), and inverses (reverse the shift of gaze). It is this group that will remain invariant under shifts in direction of gaze (O'Regan and Noë 2001).

Once again, then: Seeing a spot *as being red*, even if it is seen only with a brief glance, is to anticipate that any scanning around that spot would yield the "red" invariance group of changes in patterns of receptions. Seeing as red is anticipative.

These examples have in common that contentful experience, from seeing something as a toy block, to a boundary as being straight, to a spot as being red, are all anticipative. They are akin to Gibson's affordances (Bickhard and Richie 1983; Gibson 1966, 1977, 1979) or Piaget's schemes, but extended to all scales of interaction.

Perception is interactive, and perceiving *as* is (interactively) anticipative. Perception is interaction in the service of the apperceptive maintenance and updating of situation knowledge. It is not the contact that is perceived, it is the content.

## 4.2 Evolutionary epistemology

It can be tempting to consider that the world impresses itself into a passive mind, like a signet ring into wax, in some sort of transductive encoding, or that the world scratches itself over time into the wax of the mind in some sort of induction. But such models violate the most basic desiderata for models of representation. Among other problems, they make the possibility of error a mystery and system detectable error impossible.

But if representation is emergent in anticipative interaction systems, not in externally related structural or causal or informational or nomological correspondences, then there is no such temptation to assume that an interactive system can be impressed by the world into an otherwise passive mind. Interactive systems must be constructed, not passively received.

And, if such constructions are not prescient, those constructions will be subject to the possibility of being in error, and must be subject to being selected out if such error is encountered. An (inter)action based model of representation, thus, forces a variation and selection constructivism. It forces an evolutionary epistemology (Campbell 1974).[31]

Much more must be developed for a general model of learning. Heuristically partially foresighted constructions, in particular, require significant further considerations (Bickhard and Campbell 1996). Some form of evolutionary epistemology, however, is the outer framework within which more sophisticated kinds of learning must function.[32] It is not possible to knowledgeably go beyond whatever knowledge you already have.

If the constructive processes involved are themselves *recursive* in the sense that they are on the grounds of, in the context of, and generated as variations of, prior constructions—if prior constructions constitute a resource for further constructions—then there emerges the possibility of enabling and constraining relationships between constructions already attained and further possible constructions. Such enabling and constraining relationships are the subject matter of development.

So, a recursive constructivism yields a natural model of development, and the recognition of the importance of interactions between constructions and environment in such development is sometimes called *interactionism*. Since interactivism forces an evolutionary epistemology, which, in turn, forces an interactionism of development, we have, in slogan form: interactivism forces interactionism.

## 4.3 Language

It is with respect to perception and language that encodingist presuppositions seem most deeply embedded in our intuitions: encodings of the world into the mind in the case of perception, and encodings of mental contents into utterances in the case of language. Encodings, however, do not work in either case. Encodings are stand-ins, and must already have prior representations to stand-in for. This cannot be the case for basic representations that cross an epistemic boundary, either from an epistemic agent to the world or from the world to an epistemic agent. *Given* such representations, *derivative* stand-in encodings can in principle be defined, but encodings cannot cross such boundaries themselves.

[31] Thus, neither transduction nor induction exist as *generators* of representations. Transduction, however, more innocently refers to changes in form of energy, and induction, not so innocently, to kinds of warrant (for already existing representations) based on "confirming" and "disconfirming" instances. See Bickhard (2002) for an account of such warrant.

[32] And it requires relatively little addition to a basic evolutionary epistemology to be able to account for, for example, classic kinds of conditioning (Bickhard 2006).

Language, therefore, cannot be constituted as encodings of mental contents. But what else could language be, if not encodings?

The basic form of interface between an epistemic agent and the world is interaction. So, it would seem, utterances should also be forms of interaction. But what kind? Interactions with what?

A first obvious candidate would be "Interactions with others' minds." Utterances induce alterations[33] in the minds of the audience. I argue that this is partially correct, but it cannot be the full story. If others' minds were the proximate locus of interaction, then it would prove at least difficult to account for the difference between someone understanding a command, for example, and their obeying it, or understanding a declarative sentence and believing it, as well as for the difference between the utterer's understandings after making an utterance and those of the audience. Simply, there is something *social* about a public utterance (or written sentence, etc.) that is missed if the focus is solely on individual audience minds.

#### *4.3.1 A social ontology*

Consider an agent apperceiving his or her environment. Many of the interactive potentialities available can be apperceived on the basis of perceptions of walls, chairs, rocks, etc. The affordances are relatively direct given the objects and their layout.

If the environment contains another agent, however, then an interesting problem emerges. The interactive potentialities of another agent are not so clear just given a simple perceptual contact. Most especially, there are important senses in which the interactive potentialities of another agent are dependent on that other agent's interactive characterization of his or her environment—and that environment includes the first agent, and, in particular, the first agent's interactive characterization of the second agent.

There is an inherent reciprocal character to the problem of characterizing a situation involving other agents. In general, there are only two possibilities for such characterizations: either one of the agents induces a characterization in the other that the first knows to be false, as in deceit, manipulation, espionage, and so on, or else the mutual participants in the situation arrive at a fixed point in which each characterization is complementarily consistent with the others. Such a mutually consistent relationship of interactive characterizations of the situation constitutes a solution to the joint problem of arriving at a stable interactive characterization of the situation. In this sense, it constitutes a solution to a coordination problem (Schelling 1963), and, therefore, a convention in a general sense related to Lewis' model of convention (Lewis 1969). I call such a joint interactive characterization a *situation convention*—a convention about what the situation is.[34]

---

[33] Not via decoding of those utterances, but via apperceptions based on those utterances.

[34] This model is like Lewis's in many respects, but also different in several crucial respects. These differences make a great deal of difference for issues of social ontology, for example, especially, though not exclusively, normative aspects of social ontology (Bickhard 2004b, forthcoming-c, in preparation). I will not develop these considerations here, but, instead, outline a few of the points about language that this model can frame and support.

### 4.3.2 Interactions with social realities

Situation conventions constitute a realm of social reality that utterances can interact with. So long as the commonality or coherence among the respective situation knowledge webs of each of the social participants is maintained over the course of their respective apperceptions of the relevant utterances, the social realities will themselves be changed, but will remain social in the sense of being situation conventions. So, utterances interact with—create, change, repair, and so on—social realities in the form of situation conventions. This can occur with respect to institutionalized forms of situation convention, such as calling a meeting to order, or with respect to non-institutionalized situation conventions, such as bringing a topic to the forefront of discussion so that, among other consequences, an appropriate pronoun will be resolved in terms of that topic.[35]

Language is not the only way in which social realities can be interacted with, but language constitutes a(n institutionalized) convention for the productive construction of utterances that have conventional interactions with situation conventions—language is constituted as a conventionalized system for interacting with conventions.

### 4.3.3 Properties and consequences

Both perception and language, then, are particular kinds of interaction. Neither is an encoding phenomenon. The fundamental nature of autonomous systems is their necessary ongoing interactions, and both perception and language are differentiations and specializations of such interactions, with language involving a social level of emergence.

The shift from utterances as encodings to utterances as interactions is fundamental, and has multiple fundamental consequences.

*Context dependency.* If utterances interact with situation conventions to produce resultant situation conventions, then the results of such interactions will, in general, depend as much on the contextual situation convention as on the utterance that interacts with it. This is similar to the sense in which the result of a mapping or function depends on the argument to which it is applied as well as on the mapping per se. Context dependence, according to this model, is ubiquitous. It is not limited to demonstratives and indexicals, as for Kaplan's character (Kaplan 1979, 1989; Perry 1993).

One theme in language development is an increasing ability to use language with respect to more and more general audiences, especially with written language. That is, to learn to minimize the context dependency of our language. Taken to a limiting case, and focusing just on words, this yields the classic paradigm of the word as Name that encodes whatever it represents. This is akin to treating all functions as constant

[35] This is non-institutional in the sense that it occurs only with respect to the participants in this situation at this time, and is likely to have never occurred before as well as to never occur again. Institutionalized conventions, in contrast, occur over ranges of people and times (Bickhard 1980, 2004b, forthcoming-c, in preparation). Note that Lewis's original model cannot address non-repeated, non-institutionalized, conventions because it is defined in terms of behavioral regularities.

functions, producing the same result no matter what the argument. This is not only a limiting case of context dependency, it is an *unreachable* limiting case for language. And, even if it is granted for the sake of argument, there are still further fundamental errors involved.

It is an unreachable limiting case because the necessary context of common assumptions about how utterances are to be understood—some commonality of language—can be diminished, but cannot be eliminated, even for the *most* context independent case of formal language. It is also unreachable because a focus on words cannot account for the meanings or consequences of full utterances. Words must integrate in some way in order to produce utterances, and a restriction to words as Names undoes Frege's introduction of operators for quantifiers[36] and encounters Russell's impossible problem of the unity of the proposition—how can single elements, even if it is accepted that they are simple encodings of particulars, combine to produce something more than another encoding (Hylton 1990)?

Even if such problems with construing words as encodings are overlooked, there are further errors. A constant function on the integers, even if strictly constant, is not itself an integer. Similarly, even if a word did differentiate a single particular in all contexts, the process by which such differentiations can occur is not the same as that which is differentiated—a unit set is not the same as its element. And, for yet one more problem, there remains the difficulty that a *differentiation* of a particular, even a unit set differentiation, is still not a *representation* of that particular. A (specification of a) unit set is not the same as (a representation of) its element.

In other words, taking the processes of interactive differentiation and transformation of situation conventions to unreachable extremes of particularity, and conflating process of differentiation with what is differentiated, yields common encodingist construals of language. This perspective highlights the errors of an encoding approach to language and shows how such an approach constitutes a degenerate form of a more realistic conception of language as a(n interactive) toolbox (Wittgenstein 1953; Bickhard 1987).

*Utterances are not representational.* A second consequence of this interactive character of utterances is that they are not themselves representational. Utterances can be used to create representations for an audience, and to make such representations available in a situation convention, but the utterances per se are interactions with (relationships among) representations, not representations per se. Again, taking an utterance as directly representational is akin to construing a constant function on the integers as being itself an integer.

And again this point is an extension of noticing that it is utterances, not sentences, that generate (representations with) truth value. Because utterances are context-dependent, the focus must be on utterances rather than sentences, and because utterances transform contextual situation conventions, that may represent the world, into consequent situation conventions, which may also represent the world, it is the

[36] Frege introduced operators into language for quantifiers, but he retained a basic encoding conception for other kinds of words, and ended up with an unfortunate hybrid of operators and encodings. Utterances as interactions extends the operator intuition to all of language.

representations involved in those situation conventions that bear truth value, if anything does, rather than the utterance interactions with them.[37] In a more limited model, it is the products of Kaplan's characters that are representations ("propositional contents", in his framework), not the characters themselves. Taking utterances to directly have truth value is akin to taking a function on the integers as being prime or non-prime.

*Syntax, semantics, pragmatics.* The division of the study of language into syntax, semantics, and pragmatics is taken to be a theory neutral differentiation within the subject matter. There can be disputes about just where the boundaries are, but little question of the basic distinctions.

In the interactive model, in contrast, the properties and phenomena of language do not aggregate in the same groups as within an encoding framework. In particular, standard conceptions of syntax, semantics and pragmatics do not apply. For example, truth value, insofar as it is involved, is a property of the resultants of some utterances, not of sentences or utterances per se. Truth value is, within the standard distinctions, a property of the meaning of sentences, and truth *conditions* are often taken to *constitute* that meaning (at least for declarative sentences). But truth value as a property of *resultants* of utterances should be a property of one of the kinds of purposes for which language can be engaged in—truth value should be a part of pragmatics. Truth value is a property that can be pragmatically aimed for. It is not inherent in what an utterance does or how it does it.

Conversely, the "pragmatic" power of an utterance *form*, the ways in which it might interact with its context, constitutes the "meaning" of that utterance form in this view. Such interactive power is the only realm of properties that can be (partially) abstracted away from usages in particular contexts.[38] But this yields "meaning" as pragmatic power, not as truth conditions.

*Universal grammar.* I will mention one further consequence of viewing language in this interactive manner, a consequence for universal syntactic constraints on language—for UG. If utterances are interactions with, operators on, situation conventions, then words are sub-operators which combine in particular ways to produce full situation convention operators. But they do not functionally combine in all possible combinatoric patterns.

Constructing an utterance out of sub-utterance types is akin to constructing a function in recursive function theory out of a base set of generating functions. In the case of language, however, sub-utterance types are differentiated out of full utterances and out of prior sub-utterance types, rather than defining a base generating set first (as is the case for recursive function theory). The most appropriate approach to describing

[37] This, of course, ignores non-declarative utterances. An interactive model addresses these even more directly: e.g., interactive transformations of goals are just as much a possible product of utterance apperceptions as are transformations of representations (Bickhard 1980).

[38] The "partially" caveat stems from the point that utterances not only do not encode particulars, they also do not encode mappings or functions or transformations. Encodings will not work in any form here. Instead, utterances must be apperceived. When that apperception goes smoothly, this process can seem algorithmic, but, in other cases, it can involve hermeneutic problem solving to try to figure out what is "meant" (Bickhard 1995).

the constraints that appear in such differentiations and their possible combinations is that of categorial grammars, in which sentences are constructed out of sub-sentence types, and the types are differentiated starting with full sentence types.[39]

When the functioning of such utterances-constructed-out-of-instances-of-sub-utterance-types is examined, interesting and powerful constraints on how utterances can be structured appear. For example, the cognitive organization in an interactive model is highly relational: various interactive possibilities conditionally indicating other possibilities. How could a transformation of such a relational organization be constructed? A first requirement is that some local realm of the organization be differentiated that can serve as that-which-is-to-be-changed, as the target of a transformation. Second, the transformation must be somehow specified. Already with just these considerations we have functionally necessary differentiation between a target and a transformation of that target—a kind of logical subject and predicate. So, some sort of subject-predicate structure is functionally forced by the nature of the problem. Note that the bare conceptual possibility of this constraint suffices to show that Fodor's argument that thought must have a subject-predicate (propositional) structure because language does is unsound (Bickhard 1995, 2007, in preparation).

More detailed consideration of how such differentiations and transformations could function yields several variants of locality constraints—everything must function locally in such a relational structure because it is only local regions that are functionally accessible at any time. These locality constraints happen to capture at least one version of UG (Bickhard 1995).

I will not present the details of how this is so because my primary point here depends solely on the very possibility of such a functional derivation: if universal grammatical constraints can be derived functionally, or if it is even conceptually possible that they can be so derived, Chomsky's poverty of the stimulus argument is shown to be unsound. His argument is that there must be constraints on the space of mathematically possible grammars in order for language learning to be possible; he then claims to eliminate the environment as a possible source of such constraint, and concludes that those constraints therefore must be innate. But, his argument is an argument by elimination, and he has at best eliminated the environment. He does not address the possibility of such functional constraints. Functional constraints too can reduce the space of mathematically possible grammars, and do so in a functional manner, whereas for Chomsky the only function of the constraints is that they make learning possible. Otherwise, they are logically and functionally arbitrary. In the interactive functional derivation of such constraints, they are not arbitrary, and, because they are not arbitrary, they do not have to in some mysterious sense be part of the "knowledge" of the language learner. The learner will confront the consequences of such constraints in failures of communication of various sorts (note that this is a type *functional* feedback that Chomsky does not address in his arguments against "environmental" sources of constraint). That is, the structure dependencies of language will be learned just like the structure dependencies of any other complex task structure (e.g., learning how to repair engines)—by an evolutionary epistemological process.

---

[39] But a categorial grammar approach cannot work in this interactive context without some significant changes (Bickhard and Campbell 1992).

Ultimately, then, grammar is a framework of constraints—functional constraints—on how to construct utterances for interacting with situation conventions. Because the constraints are functional, and functionality is itself context-dependent and graded, they will emerge and be used in context-dependent and graded manners. In consequence, there is no clean cut between syntax and meaning (Goldberg et al. 2007). Various other kinds of functional considerations, such as processing demands and statistical properties of exposure, will be functionally relevant for what is classically taken to be in the realm of austere mathematical syntax (Diessel 2007; Hawkins 2007). Even sound differentiations will be functionally graded, not foundational (Port 2007). And functional considerations will include the social and emotional aspects of social interaction as well as, or even rather than, the formal aspects (Greenspan and Shanker 2007).

Language and language learning are not formal party games of learning abstract structures. Language and language learning are phenomena of social and emotional interaction, with a special (conventional) productive tool kit for constructing such interactions and with an emergent realm of social realities as foci of interactions. There are important structural constraints on how such interactions can function, and languages and language use will tend to honor such constraints insofar as and to the extent to which they are functionally relevant—not to the extent to which they are prescriptive, especially not innately prescriptive.

#### *4.3.4 Some examples of language processing: differentiation and context dependency*

I will illustrate some of these properties of the processing of linguistic interactions with a focus on the sentence:

> The man who gave his paycheck to his wife was wiser than the man who gave it to his mistress. (Partee 1972).

Primarily, I will be illustrating various forms of differentiation and context sensitivity. Partee's sentence is quite nice in this respect in that it involves a double context sensitivity.

In particular, the pronoun "it" is not co-referential with its antecedent. The antecedent is "his paycheck", but this is already in itself a context sensitive differentiator. In the context of the "it", it differentiates the second man's paycheck, not the first man's paycheck.

In more detail, the "it" is a context sensitive evocation of its antecedent. In this case, the antecedent is itself also context dependent—"his paycheck"—but now it is evoked in a different context, so that it differentiates and evokes a representation of a paycheck associated with the second man. The pronoun is not co-referential with its antecedent—instead, it is "co-context-sensitively differentiating" with its antecedent. In general, reference, like truth value, is a pragmatic goal that *might* be attained, using various context sensitive tools in the language tool kit in particular contextual situations. Neither are inherent properties of utterances or sentences per se.

In fact, such non-coreferentiality is not uncommon, and, even more generally, pronouns don't necessarily require any reference at all in order to function correctly—they can differentiate *indefinite* members of a class, perhaps even *different* indefinite members of a class. Contrast:

John lost *a black pen* yesterday and Bill found *it* today.

With:

*My home* was once in Maryland, but now *it's* in Los Angeles.
John thinks *my home* is in Maryland, but Bill thinks *it's* in Los Angeles.
We need *a secretary* and we need *her* soon.
John couldn't catch *a fish* if *it* jumped into his lap.
(Partee 1972)

Functioning via differentiation is not limited to pronouns. The opening of the sentence, "the man", differentiates—evokes—a representation of a man. It is a differentiation in the sense that the word apperceptively evokes the *kind* of representation appropriate to the word "man", as differentiated from other kinds of representation.

"The" evokes a claim that there is some particular man, available in some sense in the current (linguistic) situation convention that is being differentiated. Without more context for the utterance of the sentence, it is not clear what man will or could be differentiated. It could be someone being previously discussed, or perhaps someone among a group of men who were being discussed. It could be that only one member (or, perhaps, none) of the audience shares a prior history with the utterer such that they will resolve this differentiation in common, and commonly understood, ways.

Words and phrases are tools. A major kind of tool use is to differentiate, and *sometimes* to attempt to achieve (singular) reference via such differentiation: "The roast beef at table three needs water."[40] Reference is a pragmatic achievement, when it is achieved. And at times it is not even part of the goal of the utterance.

And this point holds even for proper names: Which "John" is being referred to? Only context can resolve the referential differentiation. Even the most classic cases, such as "The Empire State Building" succeed in differentiating a particular in almost all contexts only insofar as it is contextually commonly (culturally) understood that this is so. This is about as broad a context dependency as is possible, but the dependency on context is still not absent.[41]

The "paycheck" sentence builds up a relational organization of representations of two men—and constructs the relation "wiser" between them—and does so in ways that are equivalently available to all members of the social situation, assuming that they equivalently share relevant past experiences together—past conversational and non-conversational experiences. Differentiations differentiate within classes of alternatives, and linguistic differentiations can function as social tools primarily within socially shared contexts that provide such classes of alternatives.[42]

---

40 For extensive discussions of this and related examples, see, e.g., Nunberg (1979) and Fauconnier (1985).

41 See Bickhard (2003a).

42 At times, a language use can create a socially shared framework *in virtue* of the use making certain presuppositions about what is shared. Even if such a presupposition is false at the time of the utterance, making that presupposition may evoke in the audience a recognition that it is being made, and thereby create

Words and phrases are differential evocators of representational apperceptions, and representations are differentiators of the world. Both are context dependent processes, thus the possibility of multiple layers of context sensitive differentiation.

Modeling language understanding in terms of processes of apperceptive differentiation and construction is a more powerful framework than a naming or encoding or denotational framework. The singularities that are assumed in such approaches can be modeled as *successful use* of tools to achieve singular differentiations (though never guaranteed successful). So phenomena that seem consistent with such encoding approaches can always also be handled by the interactive model.

There are many examples, however, including a few listed above, that are at best difficult to account for with(in) an encodingist, Naming, framework—but that are nevertheless *natural* manifestations of apperceptive differentiation and construction, and of the context sensitivities that those entail. Context sensitivity is ubiquitous, not limited to just demonstratives and indexicals. And context sensitivity requires modeling in terms of some sort of "taking into account" of the context to produce the consequences of utterances—it requires interactions with (social) contexts to produce consequent (social) contexts.

Most basically, encodingism is logically incoherent and impossible. This holds just as much for language as for any other kind of phenomena. The impossibility of encodingism together with ubiquitous context sensitivity *forces* some kind of interactive approach.[43]

## 5 Conclusion

Issues of representation permeate everywhere in psychological and social phenomena—perception, memory, motivation, learning, emotions, consciousness, language, social ontology, and so on and on—and issues stemming from metaphysical presuppositions of substance versus process extend even further (process metaphysics, processes and boundaries in evolution, rigidity in psychopathology, etc.—topics mostly not explored here). There are good reasons to abandon substance and particle metaphysical frameworks, and to explore process frameworks. A part of that exploration, which has its own strong rationales, is to explore interactive models of representation and cognition, rather than encodingist models. It is to explore the *multiple* kinds of emergent normativity that permeate mentality and social personhood.

We can address the same phenomena in such explorations as before, but they do not look the same. Substance and encoding presuppositions have permeated Western thought for millennia, and attaining a fresh process and interactive view is not easy. But substance and encodingist frameworks are conceptually and empirically bankrupt, and it is time for the process alternative that has been submerged so long to emerge.

Footnote 42 continued
the shared reality being presupposed. This kind of process constitutes a general framework for modeling implicature. See Bickhard (1980).

43 Functional grammar, speech act theory, conversational analysis, and others, explore important ranges of action and interaction properties of language. But they are all based on underlying assumptions of propositional encodings, and, therefore, do not avoid basic encodingist incoherencies.

Most sciences have already moved to process frameworks, some of them centuries ago. But the introduction of process and action as the framework for understanding mentality was introduced only a little over a century ago by Peirce (Joas 1993), and there is still a great deal of exploration to be done. It's about time.

**Acknowledgements** Thanks to Richard Campbell, Cliff Hooker, Alex Levine, Susan Schneider, and Johanna Seibt for very helpful comments on earlier drafts.

## References

Aitchison, I. J. R. (1985). Nothing's plenty: The vacuum in modern quantum field theory. *Contemporary Physics, 26*(4), 333–391.

Aitchison, I. J. R., & Hey, A. J. G. (1989). *Gauge theories in particle physics*. Bristol, England: Adam Hilger.

Allen, J. (2007). Stepping off the pendulum: Why only a thoroughly action based approach can fully transcend the nativist-empiricist epicycles and ground mind in the natural world. Interactivist Summer Institute, The American University of Paris, May 29, 2007.

Belsey, A. (1995). Matter. In T. Honderich (Ed.), *The Oxford companion to philosophy* (p. 539). Oxford: Oxford University Press.

Bickhard, M. H. (1980). *Cognition, convention, and communication*. New York: Praeger Publishers.

Bickhard, M. H. (1987). The social nature of the functional nature of language. In M. Hickmann (Ed.), *Social and functional approaches to language and thought* (pp. 39–65). New York: Academic.

Bickhard, M. H. (1993). Representational content in humans and machines. *Journal of Experimental and Theoretical Artificial Intelligence, 5*, 285–333.

Bickhard, M. H. (1995). Intrinsic constraints on language: Grammar and hermeneutics. *Journal of Pragmatics, 23*, 541–554.

Bickhard, M. H. (2000a). Information and representation in autonomous agents. *Journal of Cognitive Systems Research, 1*, 65–75. http://www.elsevier.nl.

Bickhard, M. H. (2000b). Motivation and emotion: An interactive process model. In R. D. Ellis & N. Newton (Eds.), *The Caldron of Consciousness* (pp. 161–178). J. Benjamins.

Bickhard, M. H. (2000c). Emergence. In P. B. Andersen, C. Emmeche, N. O. Finnemann, & P. V. Christiansen (Eds.), *Downward causation* (pp. 322–348). Aarhus, Denmark: University of Aarhus Press.

Bickhard, M. H. (2002). Critical principles: On the negative side of rationality. *New Ideas in Psychology, 20*, 1–34.

Bickhard, M. H. (2003a). Some notes on internal and external relations and representation. *Consciousness & Emotion, 4*(1), 101–110.

Bickhard, M. H. (2003b). An integration of motivation and cognition. In L. Smith, C. Rogers, & P. Tomlinson (Eds.), *Development and motivation: Joint perspectives* (pp. 41–56). Leicester: British Psychological Society, Monograph Series II.

Bickhard, M. H. (2003c). Variations in variation and selection: The ubiquity of the variation-and-selective retention ratchet in emergent organizational complexity, Part II: Quantum field theory. *Foundations of Science, 8*(3), 283–293.

Bickhard, M. H. (2004a). Process and emergence: Normative function and representation. *Axiomathes–An International Journal in Ontology and Cognitive Systems, 14*, 135–169. Reprinted from: Bickhard, M. H. (2003). Process and emergence: Normative function and representation. In J. Seibt (Ed.), *Process theories: Cross disciplinary studies in dynamic categories* (pp. 121–155). Dordrecht: Kluwer Academic.

Bickhard, M. H. (2004b). The social ontology of persons. In J. I. M. Carpendale & U. Muller (Eds.), *Social interaction and the development of knowledge* (pp. 111–132). Mahwah, NJ: Erlbaum.

Bickhard, M. H. (2005). Consciousness and reflective consciousness. *Philosophical Psychology, 18*(2), 205–218.

Bickhard, M. H. (2006). Developmental normativity and normative development. In L. Smith & J. Voneche (Eds.), *Norms in human development* (pp. 57–76). Cambridge: Cambridge University Press.

Bickhard, M. H. (2007). Language as an interaction system. *New Ideas in Psychology, 25*(2), 171–187.

Bickhard, M. H. (forthcoming-a). Interactivism. In P. Calvo & J. Symons (Eds.), *Routledge companion to the philosophy of psychology*. London: Routledge.

Bickhard, M. H. (forthcoming-b). Interactive knowing: The metaphysics of intentionality. In R. Poli, J. Seibt, & J. Symons (Eds.), *Theory and applications of ontology*. Dordrecht: Kluwer
Bickhard, M. H. (forthcoming-c). Social ontology as convention. *Topoi*.
Bickhard, M. H. (in preparation). *The whole person: Toward a naturalism of persons—contributions to an ontological psychology*.
Bickhard, M. H., & Campbell, R. L. (1992). Some foundational questions concerning language studies: With a focus on categorial grammars and model theoretic possible worlds semantics. *Journal of Pragmatics, 17(5/6)*, 401–433.
Bickhard, M. H., & Campbell, R. L. (1996). Topologies of learning and development. *New Ideas in Psychology, 14*(2), 111–156.
Bickhard, M. H., Richie, D. M. (1983). *On the nature of representation: A case study of James Gibson's theory of perception*. New York: Praeger Publishers.
Bickhard, M. H., & Terveen, L. (1995). *Foundational issues in artificial intelligence and cognitive science: Impasse and solution.* Elsevier Scientific.
Brown, H. R., & Harré, R. (1988). *Philosophical foundations of quantum field theory*. Oxford: Oxford University Press.
Butchvarov, P. (1999). Substance. In R. Audi (Ed.), *The Cambridge dictionary or philosophy* (p. 887). Cambridge: Cambridge University Press.
Campbell, D. T. (1974). Evolutionary Epistemology. In P. A. Schilpp (Ed.), *The philosophy of Karl Popper* (pp. 413–463). LaSalle, IL: Open Court.
Campbell, D. T. (1990). Levels of organization, downward causation, and the selection-theory approach to evolutionary epistemology. In G. Greenberg & E. Tobach (Eds.), *Theories of the evolution of knowing* (pp. 1–17). Hillsdale, NJ: Erlbaum.
Campbell, R. J. (1992). *Truth and historicity*. Oxford.
Campbell, R. J., & Bickhard, M. H. (in preparation). Physicalism, emergence, and downward causation.
Campbell, R. L., & Bickhard, M. H. (1986). *Knowing levels and developmental stages*. Contributions to Human Development. Basel, Switzerland: Karger.
Cao, T. Y. (1999). *Conceptual foundations of quantum field theory*. Cambridge: Cambridge, Press.
Carlson, N. R. (2000). *Physiology of behavior* (7th ed.). Boston: Allyn and Bacon.
Chang, C. C., & Keisler, H. J. (1990). *Model theory*. North Holland.
Chapman, M. (1988). *Constructive evolution: Origins and development of Piaget's thought*. Cambridge: Cambridge University Press.
Christensen, W. D., & Bickhard, M. H. (2002). The process dynamics of normative function. *Monist, 85*(1), 3–28.
Christopher, J. C., & Bickhard, M. H. (2007). Culture, self and identity: Interactivist contributions to a metatheory for cultural psychology. *Culture & Psychology, 13*(3), 259–295.
Clifton, R. (1996). *Perspectives on quantum reality*. Kluwer Academic.
Cummins, R. (1996). *Representations, targets, and attitudes*. Cambridge, MA: MIT Press.
Davies, P. C. W. (1984). Particles do not exist. In S. M. Christensen (Ed.), *Quantum theory of gravity* (pp. 66–77). Bristol, England: Adam Hilger.
Diessel, H. (2007). Frequency effects in language acquisition, language use, and diachronic change. *New Ideas in Psychology, 25*(2), 108–127.
Doyle, J. (1985). Circumscription and implicit definability. *Journal of Automated Reasoning, 1*, 391–405.
Dretske, F. I. (1988). *Explaining behavior*. Cambridge, MA: MIT Press.
Fauconnier, G. (1985). *Mental spaces*. Cambridge, MA: MIT Press.
Fodor, J. A. (1981). The present status of the innateness controversy. In J. Fodor (Ed.), *Representations* (pp. 257–316). Cambridge: MIT Press.
Fodor, J. A. (1987). A situated grandmother? *Mind and Language, 2*, 64–81.
Fodor, J. A. (1990a). *A theory of content and other essays*. Cambridge, MA: MIT Press.
Fodor, J. A. (1990b). Information and representation. In P. P. Hanson (Ed.), *Information, language, and cognition* (pp. 175–190). Vancouver: University of British Columbia Press.
Fodor, J. A. (1991). Replies. In B. Loewer & G. Rey (Eds.), *Meaninig in mind: Fodor and his critics* (pp. 255–319). Oxford: Blackwell.
Fodor, J. A. (1998). *Concepts: Where cognitive science went wrong*. Oxford.
Fodor, J. A. (2003). *Hume variations*. Oxford.
Fodor, J. A., & Pylyshyn, Z. (1981). How direct is visual perception?: Some reflections on Gibson's ecological approach. *Cognition, 9*, 139–196.

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.
Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting and knowing* (pp. 67–82). Hillsdale, NJ: Erlbaum.
Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
Gibson, R. F. (2004). Quine's behaviorism cum empiricism. In R. F. Gibson (Ed.), *The Cambridge companion to quine* (pp. 181–199). Cambridge: Cambridge University Press.
Gill, M-L. (1989). *Aristotle on substance*. Princeton.
Glock, H.-J. (1996). Necessity and normativity. In H. Sluga & D. G. Stern (Eds.), *The Cambridge companion to wittgenstein* (pp. 198–225). Cambridge: Cambridge University Press.
Glock, H.-J. (2003). *Quine and Davidson on language, thought, and reality*. Cambridge.
Goldberg, A. E., Casenhiser, D., & White, T. R. (2007). Constructions as categories of language. *New Ideas in Psychology, 25*(2), 70–86.
Graham, D. W. (1997). Heraclitus' criticism of ionian philosophy. In C. C. W. Taylor (Ed.), *Oxford studies in ancient philosophy* (Vol. XV, pp. 1–50). Oxford: Oxford University Press.
Graham, D. W. (2006). *Explaining the cosmos*. Princeton, NJ: Princeton University Press.
Greenspan, S., & Shanker, S. (2007). The developmental pathways leading to pattern recognition, joint attention, language, and cognition. *New Ideas in Psychology, 25*(2), 128–142.
Guthrie, W. K. C. (1965). *A history of greek philosophy II: The presocratic tradition from parmenides to democritus*. Cambridge: Cambridge University Press.
Hale, B., & Wright, C. (2000). Implicit definition and the a priori. In P. Boghossian & C. Peacocke (Eds.), *New essays on the a priori* (pp. 286–319). Oxford: Oxford University Press.
Halvorson, H., & Clifton, R. (2002). No place for particles in relativistic quantum theories? *Philosophy of Science, 69*(1), 1–28.
Hawkins, J. A. (2007). Processing typology and why psychologists need to know about it. *New Ideas in Psychology, 25*(2), 87–107.
Hilbert, D. (1971). *The foundations of geometry*. La Salle: Open Court.
Huggett, N. (2000). Philosophical foundations of quantum field theory. *The British Journal for the Philosophy of Science, 51*(supplement), 617–637.
Huggett, N., & Weingard, R. (1996). Critical review: Paul Teller's interpretive introduction to quantum field theory. *Philosophy of Science, 63*, 302–314.
Hume, D. (1978). *A treatise of human nature*. Index by L. A. Selby-Bigge; Notes by P. H. Nidditch. Oxford.
Hylton, P. (1990). *Russell, idealism, and the emergence of analytic philosophy*. Oxford: Oxford University Press.
Joas, H. (1993). American pragmatism and German thought: A history of misunderstandings. In H. Joas (Ed.), *Pragmatism and social theory* (pp. 94–121). Chicago: University of Chicago Press.
Kaplan, D. (1979). On the logic of demonstratives. In P. French, T. Uehling, Jr., & H. Wettstein (Eds.), *Contemporary perspectives in the philosophy of language* (pp. 401–412). Minneapolis: Minnesota University Press.
Kaplan, D. (1989). Demonstratives: An essay on semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals. In J. Allmog, J. Perry, & H. Wettstein (Eds.), *Themes from Kaplan* (pp. 481–563). Oxford University Press.
Kim, J. (1991). Epiphenomenal and supervenient causation. In D. M. Rosenthal (Ed.), *The nature of mind* (pp. 257–265). Oxford University Press.
Kim, J. (1993). *Supervenience and mind*. Cambridge: Cambridge University Press.
Kim, J. (1998). *Mind in a physical world*. MIT.
Kim, J. (2005). *Physicalism, or something near enough*. Princeton, NJ: Princeton University Press.
Kneale, W., & Kneale, M. (1986). *The development of logic*. Oxford: Clarendon.
Kolaitis, Ph. G. (1990). Implicit definability on finite structures and unambiguous computations. In *Proc. 5th IEEE LICS* (pp. 168–180).
Kuhlmann, M., Lyre, H., & Wayne, A. (2002). *Ontological aspects of quantum field theory*. River Edge, NJ: World Scientific.
Levine, A., & Bickhard, M. H. (1999). Concepts: Where fodor went wrong. *Philosophical Psychology, 12*(1), 5–23.
Lewis, D. K. (1969). *Convention*. Cambridge, MA: Harvard University Press.
McLaughlin, B. P. (1992). The rise and fall of British emergentism. In A. Beckermann, H. Flohr, & J. Kim (Eds.), *Emergence or reduction? Essays on the prospects of nonreductive physicalism* (pp. 49–93). Berlin: Walter de Gruyter.

McLaughlin, B. P., & Bennett, K. (2005). Supervenience. *Stanford encyclopedia of philosophy*. http://plato.stanford.edu/entries/supervenience/.
Millikan, R. G. (1984). *Language, thought, and other biological categories*. Cambridge, MA: MIT Press.
Millikan, R. G. (1993). *White queen psychology and other essays for alice*. Cambridge, MA: MIT Press.
Nunberg, G. (1979). The non-uniqueness of semantic solutions. *Linguistics and Philosophy, 3*(2), 143–184.
O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24*(5), 939–1011.
Orenstein, A. (2002). *W. V. Quine*. Princeton.
Otero, M. H. (1970). Gergonne on implicit definition. *Philosophy and Phenomenological Research, 30*(4), 596–599.
Owens, J. (1978). *The Doctrine of being in the Aristotelian metaphysics* (3rd ed.). Toronto: Pontifical Institute of Medieval Studies.
Partee, B. (1972). Opacity, coreference, and pronouns. In D. Davidson & G. Harman (Eds.), *Semantics of natural language* (pp. 415–441). Dordrecht: Reidel.
Perry, J. (1993). *The problem of the essential indexical*. Oxford.
Piaget, J. (1954). *The construction of reality in the child*. New York: Basic.
Piaget, J. (1970). *Genetic epistemology.* New York: Columbia.
Piaget, J. (2001). *Studies in reflecting abstraction*. (Ed. and transl. by Robert L. Campbell). Hove, England: Psychology Press.
Piattelli-Palmarini, M. (1980). *Language and learning*. Cambridge: Harvard University Press.
Popkin, R. (2003). *The history of scepticism*. Oxford: Oxford University Press.
Popkin, R. H., & Stroll, A. (2002). *Skeptical philosophy for everyone*. Amherst, NY: Prometheus Books.
Port, R. (2007). How are words stored in memory? Beyond phones and phonemes. *New Ideas in Psychology, 25*(2), 143–170.
Quine, W. V. O. (1966). Implicit definition sustained. In W. V. O. Quine (Ed.), *The ways of paradox* (pp. 195–198). New York: Random House.
Reale, G. (1987). *From the origins to socrates*. Albany: State University of New York Press.
Rescher, N. (1980). *Scepticism*. Totowa, NJ: Rowman and Littlefield.
Robinson, H. (2004). Substance. *Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/substance/.
Ryder, L. H. (1985). *Quantum field theory*. Cambridge.
Schelling, T. C. (1963). *The strategy of conflict*. New York: Oxford University Press.
Schurz, G. (1997). *The is-ought problem: An investigation in philosophical logic* (Trends in Logic, V. 1). Kluwer Academic.
Sciama, D. W. (1991). The physical significance of the vacuum state of a quantum field. In S. Saunders & H. R. Brown (Eds.), *The philosophy of vacuum* (pp. 137–158), Oxford: Clarendon.
Shapiro, S. (1991). *Foundations without foundationalism*. Oxford.
Shapiro, S. (1997). *Philosophy of mathematics: Structure and ontology*. Oxford.
Shapiro, S. (2005a). Higher order logic. In S. Shapiro (Ed.), *The Oxford handbook of philosophy of mathematics and logic* (pp. 751–780). Oxford: Oxford University Press.
Shapiro, S. (2005b). *The Oxford handbook of philosophy of mathematics and logic*. Oxford: Oxford University Press.
Smith, B. C. (1987). *The correspondence continuum*. Stanford, CA: Center for the Study of Language and Information, CSLI-87-71.
Taylor, C. C. W. (1997). Anaxagoras and the atomists. In C. C. W. Taylor (Ed.), *From the beginning to Plato* (pp. 208–243). Routledge.
Teller, P. (1992). A contemporary look at emergence. In A. Beckermann, H. Flohr, & J. Kim (Eds.), *Emergence or reduction? essays on the prospects of nonreductive physicalism* (pp. 139–153). Berlin: Walter de Gruyter.
Teller, P. (1996). Wave and particle concepts in quantum field theory. In R. Clifton (Ed.), *Perspectives on quantum reality* (pp. 143–154). Dordrecht: Kluwer.
Trusted, J. (1999). *The mystery of matter*. New York: St. Martin's Press.
Weinberg, S. (1977). The search for unity, notes for a history of quantum field theory. *Daedalus, 106*(4), 17–35.
Weinberg, S. (1995). *The quantum theory of fields. Foundations* (Vol. 1). Cambridge: Cambridge University Press.
Weintraub, R. (1997). *The sceptical challenge*. London: Routledge.

Wittgenstein, Ludwig (1953). *Philosophical investigations.* Cambridge: Basil Blackwell.
Wright, M. R. (1997). Empedocles. In C. C. W. Taylor (Ed.), *From the beginning to Plato* (pp. 175–207). Routledge.
Zee, A. (2003). *Quantum field theory in a nutshell*. Princeton: Princeton University Press.