Levels of Representationality

Mark H. Bickhard

Mark H. Bickhard
Department of Psychology
17 Memorial Drive East
Bethlehem, PA 18015
MHB0@LEHIGH.EDU

Levels of Representationality

Abstract

The dominant assumptions — throughout contemporary philosophy, psychology, cognitive science, and artificial intelligence — about the ontology underlying intentionality, and its core of representationality, is that of encodings — some sort of informational or correspondence or covariation relationship between the represented and its representation that constitutes that representational relationship. There are many disagreements concerning details and implementations, and even some suggestions about possible alternative ontologies, such as connectionism (though none that escape what I argue is the fundamental flaw in these dominant approaches). One assumption that seems to be held by all, however, usually without explication or defense, is that there is *one* singular underlying ontology to representationality. In this paper, I argue that there are in fact quite a number of ontologies that manifest representationality — levels of representationality — and that *none* of them are the standard "manipulations of encoded symbols" ontology, nor any other variation on the informational approach to representation.

Levels of Representationality

Phenomena of intentionality, with the "aboutness" of representation at its core, pose the primary historically remaining problem for naturalistic accounts of the world. They also, arguably, pose the deepest and most difficult problems that naturalism has had to face — more so than the problems of the nature of life, for example. Meanwhile, the aspirations of cognitive science and artificial intelligence to be able to understand and to construct intentional systems have broadened these concerns enormously. Intentionality and representation are now focal concerns at levels ranging from philosophy to psychology and biology, and even to engineering with respect to computers, connectionism and neural nets, and robotics (Glymour, 1988).

Intentionality, like all phenomena, involves issues of epistemology and of ontology, and of the relationships between them. On the epistemological side, there are the behaviors of animals that seem to involve, in some way, beliefs and desires; there are the special evidences and manifestations of language for human animals (and a few others; Savage-Rumbaugh et al, 1993); there are the internal constraints of the physiology of central nervous systems; and there is the utterly unique internal perspective and constraint of each individual's phenomenology. No other subject matter in science involves such a double perspective *internal* to the system of concern; intentionality, however, involves both physiology and phenomenology.

On the ontological side — the dual of the epistemological considerations — conceptions of the ontology of intentionality are prerequisite to any modeling or constructing interests in intentional phenomena. Investigation of the cognitive functioning of the brain, equally with aspirations of designing and building intentional systems, can only proceed within some framework of assumptions concerning the ontological nature of the intentional phenomena of interest. Unfortunately, dominant notions concerning that ontological nature of mental phenomena do not remain fixed and uncontroversial. From associations to information processing to connectionist nets, notions of the basic nature of mental phenomena have evolved and conflicted, and brain studies and design aspirations have had to both keep up with and participate in the issues and the controversies.

There was a period of one or two decades, in the 1960s and 70s, in which the information processing approach was so dominant and unchallenged that it became taken for granted by many researchers in related fields — it appeared that such foundational issues concerning the nature of intentionality could be presumed and not examined. That is no longer the case, and hasn't been since about 1980. Neither cognitive neuroscientists nor cognitive psychologists nor artificial intelligence researchers can avoid taking some sort of stand in the middle of competing alternative conceptions of the architecture of intentionality (Bickhard, 1991d). In this paper, I intend to add to the alternatives available, but in a way that is intended to offer resolutions of many of the issues and controversies involved.

There is relatively little controversy concerning issues of the epistemology of intentionality — of the manifestations of and the sources of constraint on intentionality — though there can be great controversy about the relevance and importance of various aspects and domains of these phenomena for the ontology of intentionality. Concerning the ontological explanations of and models for intentionality, however, there are vast differences. The ontological level, of course, is precisely the primary focus of concern at all levels of philosophy, science, and engineering. The ontology of intentionality is the key to the nature of intentionality, to the human and animal instances of intentionality, and to the construction of intentional artifacts. The differences and controversies permeating those issues of intentional ontology, then, have broad and central relevancies.

Dennett. Claims about design constraints on intentional ontology encompass much of the differences and controversies in the domain. Many of these constraints are derived from manifestations of intentionality, from the epistemology of intentionality. Dennett, for example, claims that intentionality is just a manner of speaking, useful for prediction of a system's behavior sometimes, and less so in other cases, but, as a manner of speaking — a stance — is equally applicable to lawn mowers, chess playing computer programs, and people, so long as it affords useful predictive power (Dennett, 1969, 1978). In this view, there would be no fact of the matter regarding the ontology of intentionality and representation. In a clarification of his position, Dennett argues that *all* intentionality is strictly derivative. In this view, the intentionality of a machine is derivative from the purposes for which humans build it and use it, and the intentionality of humans and other animals, in turn, is derivative from the purposes for which various behaviors and system organizations have been selected by evolution (Dennett, 1987). However, whatever the relationship is between evolution and the intentionality of animals, it is nothing like the ascriptive or interpretive relationship between humans and the derivative intentionality of, say, written text or a chess playing mahine. Dennett conflates these badly. This position also curiously bypasses the phenomenological perspective on intentionality. Dennett (1991) does address consciousness, but primarily in a negative way — to defeat some rampant Cartesian presumptions about consciousness — and relies on Dennett (1987) for an account of representational content, and, thus, the core of intentionality, anyway.

Fodor. Proceeding from a quite different perspective, Fodor (1975, 1981, 1983, 1986, 1987, 1990; Loewer & Rey, 1991) argues that the intentional folk psychology of beliefs and desires is not only our most epistemologically important manifestation of intentional phenomena, but that it imposes severe constraints on any intentional ontology — severe to the point of essentially reifying various aspects of folk psychology into scientifically necessary intentional ontologies. For example, he argues that the systemic variations that are possible between propositional attitudes and the propositions about which they may be held — one can believe, wish, fear, etc. some proposition P, for one form of systematic variability, and one can believe P or Q or R or ... etc., for a dual form of such variability — requires that the attitudes and the representations of what they are attitudes about must be concluded to be ontologically distinct. The attitudes end up as relatively unspecified relationships to the propositional representations. In this argument, the variability between the attitudes and their propositions yields the conclusion that they must be ontologically distinct.

Fodor makes a further application of this same form of argument to the alleged propositional representations themselves. In particular, there is an additional systematicity regarding the propositional representations from which Fodor claims a similar ontologically reifying constraint. Any intentional system, so the argument goes, that can entertain "John loves Mary" (can believe, wish, or fear it, and so on) can, in principle, also entertain "Mary loves John". Just as the systemic independent (relatively) variability between propositional attitudes and their propositions leads Fodor to claim that the attitudes and their propositions must be ontologically distinct elements or states, so does the systemic variability of various aspects or *components* of propositions yield, for Fodor, the conclusion that those *within*-propositional variabilities correspond to ontologically real elements, *sub*-propositional elements. These will be subpropositional representational elements, therefore intentional elements. This, of course, underlies the familiar claim that there is, and must be, an (innate) language of thought in the brain to be able to account for manifest thought and language (Fodor, 1981; for fundamental criticisms, see Campbell & Bickhard, 1987; Bickhard, 1991b, 1991c).

Note, however, that the very form of argument here is invalid. Systematic variabilities, even if accepted fully and without caveat, license at best a conclusion of relatively independent degrees of freedom in the space of possible constructions of whatever manifests those variabilities — be they thoughts or utterances or anything else. One way to instance, and, thus, one way to explain, such independence of constructive freedom is in terms of independent ontological atoms with broad combinatorial possibilities and constraints. The constructions, in this view, are simply the combinings of those elements or atoms, and the variabilities follow from the freedom of choices available in the course of such constructions. This, clearly, is Fodor's choice. Some version of such representational atomism is also the massively dominant choice in philosophy, in cognitive science, and in artificial intelligence.

But this is not the only possibility. *Any* such freedom of construction will manifest similar systemic variabilities. For example, constructions of chords out of pure sine waves will manifest unbounded such freedom and potential for variability. Constructions of computer programs out of subroutines or objects or agents will similarly manifest potentialities and constraints of systemic variability and substitution. Constructions of actions out of subactions, or, better, construction of a theme of action (say for a dance, out of aspectual themes, such as smoothness, rapidity, volume, freneticness, repetition, and so on) is still another instance of systemic variability and constraint. This point does not settle anything about representation and intentionality, but it does demonstrate that accepting the variabilities that Fodor claims does not necessarily *commit* to the representational atomism that he concludes. I argue, in fact, that Fodor's representational atomic encodingism, and, similarly, that of the dominant positions in philosophy, cognitive science, and artificial intelligence, are not only not necessary, but that they are logically impossible — they are at root logically incoherent (Bickhard, 1992a, 1993; Bickhard & Terveen, 1995). So, since those systematicities seem to be well supported, it's a good thing that those systematic variabilities do not have to be denied in order to escape the atomism that they seem to support!

The Cartesian Gulf. There is a different assumption, in common to Dennett and Fodor, however — and to virtually all other approaches to these phenomena — that I wish to take as the focus of this paper. This assumption is a vestige of Cartesianism, but one less noted than others. This is the assumption that representational intentionality is ontologically unitary, that it is some sort of singular natural kind — that there is some sort of singular gulf between the ontology of the mental and that of the non-mental.

I wish to argue, to the contrary, that there are at least several different sorts of representationality, and that they involve differing underlying ontologies. From this perspective, the assumption that intentional representationality all involves some kind of common underlying nature is *part* of the explanation for how such widely different positions can be held in the literature: everyone is looking at different parts of the elephant and reifying his or her part to the whole beast.

I will be approaching this discussion in a somewhat unorthodox manner. In particular, instead of marshaling various points about the epistemology of intentionality and trying to show that they constrain the ontology of intentionality in some way or another (for a discussion of some of these strategies, see Shanon, 1991; note that this methodology *presupposes* that there is a common underlying ontology), I will primarily be examining various aspects and possibilities of an ontological model (actually a partially hierarchical series of ontologies) and arguing that those aspects and possibilities would — and do — all manifest themselves in ways that get commonly described in folk psychological intentional terms. What I end up with is not a disjointed heap of intentionally described phenomena, but more like a model of various levels of representationality, with interesting interrelationships and dependencies among the levels. I begin with some very implicit and aspectual representational characteristics of systems, and end with fully differentiated and specialized representations. Along the way, I connect with various other claims about representational intentionality in the literature, universally with the conclusion that those claims have captured something correct about the manifestations of the phenomena, but that those claims are fundamentally in error concerning the underlying ontologies for those phenomena.

Levels of Representationality

Level 1 Pure intentional stance

Level 2 Functional presupposition

Level 3 Interactive implicit definition
- Differentiation

Level 4 Functional interactive predication

Level 5 Procedural implicit definition
- Iterations of presuppositions
- Unboundedness of implicitness

Level 6 Situation image
- Apperception

Level 7 Objects
- Memory
- Perception

Level 8 Learning
- Off-line processing
- Strategies
- Internal encodings
- Differentiation and defeasible reference

Level 9 Language
- Situation conventions
- Utterances as operators

Level 10 Awareness
- Consciousness as reflexive awareness
- Knowing levels and developmental stages

Levels and Levels

Level 1. I will take level 1 of representationality to be the pure intentional stance, applicable to lawn mowers and humans alike, with no particular ontological commitments or claims. This in spite of Dennett's claiming that this is not what he had in mind (Dennett, 1987 — see Dennett, 1969 and 1978, however). Such a null or at least minimal ontological position is an appropriate initial level for my purposes since the hierarchy of levels that I propose constitutes a progression of increasingly differentiated and specialized representational ontologies, and a null ontology is a convenient null point from which to begin such a hierarchy.

One characteristic that will *not* hold for this hierarchy is that it is not, and, I would argue, cannot be, a hierarchy of emergence or evolution or design of intentionality or representationality. It certainly will have relevance for the design problem, for example, in that the various ontologies presented constitute alternative approaches, and in some cases arguably necessary approaches, for instantiating various kinds of representational ascription. But, because representational ascriptions all tend to have the same canonical form of belief-desire folk psychology, it is not always easy to determine from a given ascription what sort of ontology might underlie it. In fact, if this model is correct, there is a systematic ambiguity for most intentional ascriptions among the several alternative intentional ontologies. The differing ontologies, however, will tend to sort themselves out with respect to their potentialities and counterfactual properties.

I should note that, for the most part, there is nothing magical or inherent about these levels. They are in several cases distinguished by differences that could for other purposes have been collapsed together; differentiating these levels depends on what sorts of differences seem worth emphasizing for the purposes at hand. This is in contrast, for example, to some hierarchies that are ontologically intrinsic and forced by the nature of the phenomena at hand (Campbell & Bickhard, 1986, 1992b). (And, there are a number of potential forms and levels that will not be developed here, such as, for example, emotions, values, rationality, the social person, and identity; Bickhard, 1980a, 1991a, 1992a, forthcoming, in preparation-b; Campbell & Bickhard, 1986.)

The intent in presenting a number of levels of the ontology of representationality is partly to outline an overall modeling approach, and partly to make a prima facie partial case that this approach *might* be competent to model *all* intentional phenomena. Insofar as that competence is plausible, it becomes relevant to contemporary approaches that there are *many* underlying ontologies in this approach to intentionality — not just one — and, further, that *none* of them have the standard form of representations as some special kind of representational correspondence — of representations as encodings.

Level 2. Level 2 is a kind of intentionality of presupposition involved in some system activity. For initial motivation, consider presupposition in utterances. One famous example is "The King of France is bald." A classic issue concerning this and related sentences is that, since France doesn't have a king, is this sentence false, or is it defective in some other way — and, if so, what other way? My focus, however, is on the relationship between the sentence and France having a king: a presuppositional relationship. Note that a presuppositional relationship is not the same as an entailment relationship: "The King of France is not bald." involves the same presupposition concerning France having a king, but not the same entailments. Instead, a presupposition is a kind of condition for the *appropriateness* of the sentence (Bach & Harnish, 1979; Bickhard, 1980b); uttering the sentence presupposes conditions under which that utterance is appropriate.

It is this appropriateness aspect that I will pick up on. To shift to a more explicitly functional perspective, consider the lowly and classical thermostat: it functionally presupposes that the heat flow in or out of its space will not exceed certain levels, otherwise it could not keep up; it also presupposes that the temperature in its space will not oscillate above and below the set point with greater than some frequency, otherwise the thermostat's switching mechanism would not be able to keep up. Neither of these presuppositions are "in" the thermostat in any ontological sense; they are not written on the wires. They are, instead, truths about the functional relationships between the thermostatic system and its environment: Truths about the conditions of good functioning of the system. They are functional presuppositions.

Attribution of such presuppositions requires first an attribution of something like purpose to the system. Functional presuppositions are the conditions under which the ways in which the system goes about attempting to achieve its goals will in fact tend to succeed in reaching those goals. Such purposive attributions could themselves be purely level 1 ascriptions, such as the presuppositions that my lawnmower apparently makes in its attempts to maximally befuddle and enrage me, or they could be based on an ascription of explicit goal directed feedback functional organizations to the system, as for the thermostat.

There is nothing metaphysically odd about functional presuppositions; they are simply aspects of the functional properties of purposive systems that interact with their environments. They are functional, *relational*, properties, that hold or do not hold much like any other properties. They are a kind of implicit property of the organization of the interactive system with respect to possible environments. There are, in fact, other versions of such functional implicitness: for example, a tendency for achieving certain conditions may constitute an implicit *goal*, as in the case of the guided missile that has an implicit goal of destroying an airplane, but whose explicit feedback systems are built around vector signals from infrared detectors. The explicit goal might be to achieve a discrepancy signal of zero between that vector signal and the major axis of the missile (clearly over-simplified, but sufficient for the point).

What's most interesting for my current purposes is that, in the case of humans (and often times other animals) when we observe behavior that seems to involve presuppositions of appropriateness, we tend to think and talk and write about those presuppositions as if they were full canonical beliefs. If we see someone being cautious, for example, we are likely to attribute a belief that there is a likelihood of danger. Such an attribution may or not hold, but in many respects that issue turns on what is taken to constitute, to ontologically constitute, beliefs in the first place. Certainly there is much animal and human behavior that involves sometimes elaborate functional presuppositions that do not involve any explicit encoded representations, nor conscious representations, nor perhaps even *potentially* conscious representations. Does the spider *believe* that food flies and can be caught in the spider's web? How about the frog *believing* that the moving dot is a fly?

Such examples do not seem to give us too much trouble — they seem fairly clearly to not be full examples of genuine belief. Just what they are examples of, however, is not so easily consensual. I am suggesting that they are functionally presuppositional, and that it is those presuppositions that are involved in the evolutionary selection of the respective mechanisms. In particular, the presuppositions of the respective mechanisms for eating tend to be *true* in that species' typical environments (Millikan, 1984).

More troublesome examples show up when we look at instances of *false* presuppositions, especially in humans. Consider John's typically reticent and cautious behavior whenever he is around women. In some sense, John functions in the world as if women were universally, perhaps inherently, dangerous. Maybe we can even postulate something about the kind of danger women present to John from various more detailed aspects of his behavior. These behaviors, and their aspects, manifest precisely the sorts of functional presuppositions under discussion. That is, they are the conditions under which John's behavior *would be* functionally appropriate.

If we ask John, however, he may respond that, on the contrary, he is very lonely and in fact is quite drawn to women, or maybe even deny our attribution vehemently. Or, after some discussion and reflection on his behavior and his past, he may agree with us that he seems to be afraid of women, but, in doing so, he is concurring in our inference about that, and still has no awareness of such fear.

In such cases, both we and John are overwhelmingly likely to attribute some sort of belief in the dangerousness of women to John, and then to puzzle over how he could believe such a thing and not even know it. That is, we will use canonical folk psychology belief-desire ascriptions, and then puzzle over the absence of one usual and even intrinsically expectable concomitant — awareness of one's beliefs. We do have a ready made solution to this puzzle, however, and that is to postulate an intentional realm in which such beliefs can reside that is somehow separate from and epistemically cut off from normal consciousness and normal reporting: the dynamic Unconscious. When we combine such ascriptions of *implicit* presuppositions with those of *implicit* goals or desires and other forms of implicitness, in fact, we account for a very large portion of the phenomena that we otherwise ascribe as explicit belief-desires, but in the Unconscious (Bickhard & Christopher, 1994). Note that we do *not* make the equivalent move for the spider or the frog or the bird feigning a broken wing and moving away from its nest.

The general point of level 2 of representationality, then, is that properties of the world can be presupposed in the functioning of a system without necessarily being explicitly represented in that system. Such functional implicitness is a real property belonging to a system, a real property of the functional relationship between the system and its environment, but is not a distinct representational ontology in the system. Such presuppositional representationality can hold of many systems, so long as we ascribe purpose of some sort, such as chess playing programs, missiles, thermostats, animals, and human beings. In most cases, such ascription will neither warrant nor require any ontology stronger than such functional presuppositions.

Level 3. In level 3 I consider a different kind of implicitness. Instead of the conditions that are implicitly presupposed in the appropriate functioning of a system, I wish to consider here a kind of representationality that is implicit in system functioning, system environmental interaction, that has already occurred. Specifically, consider a subsystem engaging in interaction with its environment: the internal course of that interaction will depend both on the organization of the subsystem and on the interactive properties of the environment being interacted with. That same subsystem interacting with a different environment might engage in quite different internal processes.

In particular, the internal state that the subsystem ends up in when its current interaction ceases will depend on the environment that it has interacted with. Some environments will yield the same such final state, while other environments will (or would) yield a quite different final state. The possible final states of such a subsystem, then, serve as classifications of the possible environments: each final state classifies all of the environments together that would yield that particular final state if interacted with.

Each possible final state will serve as a differentiation of its class of environments. Note, however, that, although the system will have functionally available information concerning what category of environment has been differentiated, and, thus, encountered, there is no information concerning anything about that environment beyond that it was just encountered and that it is not the same as those environments differentiated by any of the other possible final states: Minimal information.

The relationship between a possible final state and the class of environments that would yield that final state in actual interaction is a kind of interactive version of implicit definition. In model theory, a logical system implicitly defines its class of models — the potential models that do satisfy that system (Kneale & Kneale, 1986; Moore, 1988); in this case, an interactive system with a specified final state implicitly defines its class of environments — the class that would yield that final state (Campbell & Bickhard, 1986, 1992a; Bickhard, 1993; Bickhard & Campbell, 1992). The relationship between a possible final state and the class of environments that would yield that final state in actual interaction is also akin to the relationship between a set of final states in an automaton and the (implicitly defined) class of input strings that that automaton could recognize. An automata theoretic recognizer, however, unlike the interactive case, has no outputs that influence further inputs, and, correspondingly, 'recognizes' input strings rather than environments that can, in interaction, generate those input strings (Bickhard, 1980b, 1993; Boolos & Jeffrey, 1989; Eilenburg, 1974; Ginzburg, 1968; Hopcroft & Ullman, 1979).

There is a kind of implicit representation involved here: an implicit representation by virtue of interactive implicit definition. But there is no representational content involved. The system has no information about the classes of environments that it implicitly defines. Strictly, the system has no information that *anything* has been differentiated or implicitly defined: no ontology specified to this point would give the system any information that it was interacting at all, nor that there was anything to be interacting with, and, therefore, certainly not any information about what was just interacted with.

Nevertheless, when we as observers of such a system discover such implicit definitional relationships — especially in such systems that we already consider to be intentional, such as complex organisms — there is a near universal tendency to ascribe representationality to them. We, as external observers of the system in its environments, do know something about the system environment, and can note various correspondences that might be created by the differentiations, the implicit definitions, of the system activities. When we do, and most especially when the interactive differentiations are passive, with no outputs, we ascribe to those internal final states a full representational capacity as encodings, generally sensory encodings, produced by the final states of sensory transducers. As observers, we note the correspondences between system internal states and environmental conditions and reify those correspondences into epistemic relationships (Bickhard, 1992a, 1993). We assume that the system represents that which it generates correspondences with.

This move is ubiquitous (Fodor, 1990; Hanson, 1990; McClelland & Rumelhart, 1986; Newell, 1980a; Rumelhart & McClelland, 1986; Smith, 1987). It is also fraught with problems, fatal problems I argue (Bickhard, 1992a, 1993, 1995b; Bickhard & Terveen, 1995). Basically, the factual existence of correspondences gives no ground for the ascription of ontological representations in and for and to a system itself: a system doesn't know anything about any such correspondences, not their existence, nor what they are with, simply by virtue of there *being* such correspondences. Trivial versions of such internal-to-external correspondences exist between a stone and its environment;

such correspondences, in fact, are ubiquitous throughout the universe, and very few of them have anything to do with representation. This holds just as strongly for transduced encoding correspondences as it does for patterns-of-activity connectionist correspondences. It is this common assumption of representation-as-correspondence that vitiates both symbol manipulation and connectionist approaches to representation (Bickhard, 1993; Bickhard & Terveen, 1995).

This problem of the plethora of correspondences is not unknown in the literature of those who wish to characterize representation as correspondence. The general strategy, in fact, is to attempt to find additional restrictions on correspondences that will succeed in picking out just those correspondences that *are* representations (Fodor, 1990; Newell, 1980a; Smith, 1987). As a first of many critiques of this strategy, note that, even if it succeeded in its own terms, it would at best *extensionally* capture the class of representations that are correspondences: it would, in itself, give no model of representational content, of what makes those particular sorts of correspondences into representations (unless such a model of content were itself the form of restriction on the class, in which case we would have "elements in correspondence are representations if they already have representational content of what they are in correspondence with" — all the important work, in such a case, would be done by the model of content, not the model of correspondence).

For a related critique, note that there *is*, in fact, a class of correspondences that constitute representations: encodings, such as Morse code. But encodings cannot be a foundational form of representation — they cannot model representational content — because they presuppose the prior existence and understanding of such representational content. In Morse code, for example, "**. . .**" encodes "**S**" in the sense that it stands-in for "**S**" and is understood as standing-in for "**S**" — that is, "**. . .**" can be *interpreted* as standing in for "**S**" by those who already know Morse code. In turn, Morse code can be learned by learning such correspondences as between "**. . .**" and " **S**", but only if "**S**" is already known, only if the representational content of the character "**S**" is already available. "**. . .**" stands in for " **S**" in the sense of borrowing whatever representational content "**S**" already has. Such encodings, then, only borrow already existing representational content; they do not *generate* representational content. Encodings, then, cannot solve epistemological problems, because these problems require accounts of the generation of new knowledge about the situation, the world in general, mathematics, and so on. Encodings are carriers of already extant representation, not the sources of new representation. Therefore, encodings cannot be the solution to the general problem of representation.

A currently focal issue regarding representations as correspondences is the problem of the very possibility of representational error (Dretske, 1981, 1988; Fodor, 1987, 1990; Loewer & Rey, 1991; Millikan, 1984, 1993). If a correspondence (of the "right" kind) exists, then the (supposed) representation exists, and it is *correct*, while if the correspondence does not exist, then the representation does not exist, and, therefore, it *cannot be incorrect*. How can a representation exist *and* be in error?

There are several proposals in the current literature attempting to address this problem (Dretske, 1988; Fodor, 1990; Millikan, 1984, 1993). I will not address them in any detail, but simply point out that, even if any of these proposals succeeded on their own terms (and that, of course, is controversial), they would still not be acceptable models of representation. The reason is that they all depend on conditions that distinguish representational correspondences from non-representational correspondences

— Fodor's asymmetric dependencies, Dretske's learning histories, Millikan's evolutionary and learning histories — that cannot in general be assessed by the organism that supposedly has the representations. It requires an *external observer*, and a fairly knowledgeable such observer, to assess such issues as asymmetric dependencies or evolutionary or learning histories. The organism cannot do it for itself. So the purported solutions work, if they do at all, only for observers, not for the intentional organisms or systems themselves.

This is unacceptable. It violates a basic constraint of naturalism: representation, and representational error cannot be dependent on external analysis and ascription on pain of making the analysis of intentionality circularly dependent on external intentional observers. It also makes impossible the organism or artifact detecting error in its own representations, and thereby makes impossible any model of error directed processes in that organism or artifact, such as goal directed behavior or learning. Representational error must be possible, and it must be detectable — not necessarily infallibly — in the system, by the system, and for the system, independent of any observer (Bickhard, 1993, in preparation-b, in preparation-c; Bickhard & Terveen, 1995).

These critiques of encodings as exhausting or grounding the nature of representation are members of a large family of critiques of correspondence approaches to representation, some of ancient provenance, some being discovered quite recently. Even individually, they defeat an encodingist approach to representation (except for those for whom there is no alternative to encodings, so they look for some way around or some way to ignore the critiques); collectively, they bury it even deeper (Bickhard, 1992a, 1993; Bickhard & Terveen, 1995).

My current focus, however, is that there *is* something ontologically real in such correspondences, and that it is typically, though falsely, rendered in canonical representational terms. What *is* real is that the system does have functionally available information that it is in final state **A**, say, rather than **B**, or others, and that such differentiating information might be useful in further system activities — might be useful for selecting among potential further system interactions. In simple systems, such as a rock, there may be no way for such information to be used, in which case it may be causally nugatory, but that doesn't alter the point that it is present. In complex interactive systems, implicitly defining differentiations of the environment might be just what is needed in order to control behavior in adaptive ways — even if there is no representational content of what has been differentiated, of what is on the other end of the correspondence.

Here, then, we have an implicit definitional form of ontological representationality. This form is almost universally mis-taken as constituting full intentional representationality, but, nevertheless, it does constitute its own level of ontological representationality, a sort of pre-representationality — one that does exist and can be useful for further forms.

Level 4. In level 4 we look at a more explicit, but still not differentiated or specialized, form of intentionally ascribable system ontology. It is this level that constitutes a minimal emergence of ontological representationality (Bickhard, 1980b, 1987, 1991c, 1992a, 1993), though that is not the primary focus here. Level 4 involves a *functional* predication of interactive properties to a system environment. It constitutes the representational core of a position dubbed *interactivism* (Bickhard, 1987, 1992a, 1993; Bickhard & Campbell, 1989, 1992; Bickhard & Richie, 1983; Bickhard & Terveen, 1995; Campbell & Bickhard, 1986, 1992a; Vuyk, 1981).

In a system that is goal directed in at least a negative feedback sense, there may be more than one way available for the system to attempt to reach its goal. Such alternative goal-attempting procedures may involve differences in the environmental conditions and circumstances in which those different procedures will tend to succeed. Where such differences exist, the system would be well served by some method of differentiating the prevailing environmental conditions, and switching to goal-attempting procedures appropriate for those conditions.

Where, then, would information for the control of switching to the "right" procedure come from? In principle, it could come from anywhere, but, in general, it will be generated by differentiations of the current environment that are already inherent in the system's interactions with that environment. In particular, the differentiations and implicit definitions of level 3 representationality can provide exactly the sort of information needed to select appropriate subsequent subsystems for continuing some system interaction in pursuit of a goal. Thus, if the current goal is **G22** and the final state of the relevant differentiating subsystem is **F87**, then procedure **P232** should be selected, while for differing final states and/or goals, differing selections of subsequent procedures may be appropriate.

A goal subsystem, then, may switch among alternative available procedures, and may do so on the basis of the final states arrived at by interactive (or passive) subsystems that differentiate the environment. The final states of differentiators, then, serve as a kind of meta-switch for the goal system: the differentiating final states switch the goal system with respect to what classes of further procedures the goal system will call on, with the ultimate selection depending on the specifics of the current goal. Putting this another way, the differentiator final states *indicate* which further procedures might be appropriate, perhaps by setting pointers, or some equivalent, and the goal system selects from among those that are so indicated the procedures that might be appropriate for the goal.

Note that this constitutes a functional use of differentiating final states that makes no assumptions that those final states carry any representational content — makes no assumptions that they are representations *for the system* at all. All that is required is that they do differentiate appropriate environmental conditions *in fact*, not that they explicitly represent those conditions.

Similarly, the notion of "goal" that is required here is not necessarily representational. Goal, in the sense required, can be as simple as functional or physical switches based on set-points: they switch to subordinate procedures if the set-point (goal) is not met, and switch out of the local organization of process (control structure) if it is met. The goal switching organizations *can* be of much greater complexity, and goals can make use of representations when available, but a representational notion of "goal" is not required (Bickhard, 1993; Bickhard & Terveen, 1995). If a representational notion of goal *were* required, this would introduce a circularity into the model: representation defined in terms of goals, which, in turn, would be representational. Strictly functional, switching, goals avoid this circularity.

The sense in which I argue that this conditional switching in the service of a goal constitutes the emergence of minimal representationality is that it constitutes the emergence of *error* — the possibility that the system might be wrong — and *error for the system itself*. That is, it provides not only the possibility of error per se, but provides the possibility that the system might *discover* that it is wrong, in a way that is strictly in and of and by and for the system itself. It provides a form of error that is strictly internal; that is not dependent on any ascriptions of intentionality from outside the system itself.

Specifically, if the system fails to reach its goal, then something was in error in the indications of further interactions for that goal, and, since that failure to reach its goal is itself an internal condition of the system, information of such failure is functionally available to the system for further processing (e.g., such as switching to a different subsystem in the service of some higher level goal, or invoking some learning procedure) (Bickhard, 1993). My critical point is the functional availability of that error information, not what the system does (or does not) do with that information.

More importantly for my current purposes, however, is the sense in which such a system organization constitutes a functional predication about the environment. It is this functional predication, in fact, that is capable of being wrong, and of being discovered to be wrong by the system. In particular, the functional indication of some procedure **P232** on the basis of an environmental differentiation (level 3) constitutes a predication to that environment of the interactive properties appropriate to that procedure: it is a predication that environments of implicitly defined type **F87** are environments of interactive type **P232**.

If a frog, for example, engages in visual scans that have internal outcomes which normally differentiate a fly (at a particular location, with such and such a velocity, etc.), then those internal outcomes can be used to initiate and control tongue flicking interactions that can yield the eating of the fly. There is no necessity for any representational content *about flies* here, only a functional or control phenomena. However, there *is* representational content in the indication in the frog that it is in a tongue flicking and eating situation. That could be in error, and the frog could discover it to be in error by engaging in the requisite interactions and finding that they fail — if someone had tossed a small pebble in front of the frog, for example. That error information, in turn, could, *if* the frog were so equipped, initiate complex learning processes that might succeed in learning to differentiate more finely, so that pebbles no longer get differentiated as tongue flicking and eating situations.

Note that the representational content here is a predication of the environment of particular interactive potentialities — tongue flicking and eating in this case. The content is *not* about flies, though its appropriate activation is evolutionarily dependent the underlying differentiations picking out flies and similar insects often enough *in fact*. That is, the correspondences must be there with flies (etc.) in order for the indications of interactive potentialities to be true, but it is not what is corresponded with that is represented.

Standard approaches, of course, take the internal states that might in fact *differentiate* flies as constituting *representations* of flies. In the interactive model, in contrast, it is fundamentally action and interaction potentialities that are represented. Because they are represented by virtue of being functionally indicated, we have here the interface of emergence of representation from functional organization. This grounding in interactive potentialities, of course, raises questions about the nature of representational content that *is* of objects, such as object representation in humans. Such content is not present in simple systems, such as paramecia, or planaria, and so on, but is present (in varying degrees and forms) in complex animals such as apes and humans. The nature of such representational content is addressed later, in level 7.

Level 5. The interactive properties being predicated of an environment in level 4 indications are *not* encoded in any explicit representational form. Instead, they are themselves *implicitly defined* by the procedure being indicated: the predicated properties are the properties that would allow that procedure to reach its proper completion, its final states. In level 3, we have implicit definitions of environmental categories, and in level

5, implicit definitions of environmental interactive properties. In level 4, we have functional predications *between* these two kinds of implicit definitions.

Level 3 and level 5 implicit definitions are similar; the difference is primarily a matter of functional role relative to the rest of the system. In both cases, a class of environments is implicitly defined — those environments that would yield particular final states, or one of a class of final states, if interacted with under the control of a particular system. The final states in a differentiator, however, are themselves functionally differentiated: each final state is differentiated from the alternative final states in its potential functional consequences for further system activity. The indicated final states for an *indicated* procedure, in contrast, are functionally collected: the indication is that some one — some unspecified one — of the indicated set of final states will be attained.

The difference, then, is in terms of how a procedure is functionally used by the overall system, and a single procedure could in principle function as a differentiator at one time and as an indicated potentiality of interaction at another time. If the final states differentiate subsequent system activity, then the procedure is being used as a differentiator, while if the final states are each equivalently terminations of the activity of the procedure and, possibly, achievements of the current goal, then the procedure is being used for an implicit predication. It is even possible that a single procedure could be indicated, together with a set of possible final states — thus functioning in a predication — and then the particular final state that that same procedure actually arrived at might differentiate further activity, thus functioning as a differentiator. As a predication, a procedure is indicated as a potentiality — the final states are indicated as possible states to attain if the procedure is executed. As a differentiator, a procedure must be or have been executed in actuality — the final states will differentiate only if actually arrived at.

In these ontologies, all representationality is constituted either in terms of one form or another of implicit definition, or in terms of the functional relationships among system organizations that involve such implicit definitions. This is quite different from standard approaches, in which, for example, sensory differentiations are construed directly as epistemic encodings for the system — sensory encoding (e.g., Carlson, 1986). The presumed epistemic relationships in these *encoding* ascriptions are explicit, not implicit: the system is supposed to know what is represented by both the environmental differentiations and the interactive predications.

Nevertheless, in observing such a system, we, again, will make canonical ascriptions of beliefs and desires. To use the frog example — on the basis of a visual differentiation of its environment, the frog predicates the availability of a tongue flicking-and-eating opportunity (implicitly defined by their respective procedures). This will tend to be described — at least initially — as seeing a fly (or worm) and believing that it could be eaten with a tongue flick. In the case of a frog, however, as soon as we have made such ascriptions, they strain our intuitions, and we would like some other way to account for what is going on.

Both the initial intuitive acceptability and the intuitive strain are results of conflating various levels of representationality. In particular, the "seeing of a fly or worm" is in fact, I suggest, the *presupposed conditions* for the proper functioning of the tongue flicking procedure — level 2! The predication of appropriate tongue flicking environmental properties is an instance of level 5. And the control over such predications, and, thus, over such tongue flicking activities, that provides some likelihood that those level 2 presupposed conditions actually hold is on the basis of the level 3 visual differentiations that yield those level 5 predications.

So, on the one hand, there is lots of representationality going on here, but, on the other hand, only the predication of the tongue flicking interactive properties is strictly internal to the frog-system itself, in the sense that the frog could possibility discover an error in such a predication. All others are matters of presupposition and implicitness — not available to the frog per se, but, perhaps, important for other explanatory purposes of an *observer* of the frog. The presupposed conditions of "fly or worm", for example, play a very important role in the evolutionary explanations of the overall frog-system (Millikan, 1984).

Note that the sense in which the level 2 functional presuppositions are the conditions under which the level 5 interactive predications will actually hold is a general relationship among these levels. So long as such functional predications do ontologically occur in some system, there will be such presupposed conditions of appropriateness of those indications-predications, and those *presupposed* environmental properties may well go far beyond any ontologically explicit representationality in the system.

A baby's crying when hurt, for example, or perhaps a baby's withdrawing when upset, will involve presuppositions about the caregivers available to the infant — presuppositions of responsiveness in the first case, and, perhaps, non-responsiveness in the second — that can far exceed the cognitive capacities of that infant. Furthermore, just as utterance presuppositions can themselves have presuppositions — e.g., "France has a king" presupposes that France is a country — so also can functional presuppositions involve further presuppositions: the infant's presupposition of unresponsiveness may involve a deeper presupposition of the incapacity, the unworthiness, of the infant to evoke the necessary caring that would yield responsiveness. Such iterated presuppositions about the self can have massive manifestations in one's life, yet most often are never explicitly recognized or understood — or changed. The point here is that they do not have to be explicitly (or unconsciously) cognized in order to be presupposed.

Nevertheless, again we find that the standard ascriptions to such situations are canonically in terms of belief and desire, and the concomitant ad hoc invention of an Unconscious in which to place those presumed explicit beliefs and desires. In the case of the infant, however, this involves attributing cognitive capacities to the infant's Unconscious that are impossible for the infant until a number of months later (Bickhard & Christopher, 1994). This should have given some serious pause to these facile ascriptions, but it doesn't seem to have done so.

In the case of adult human beings, such presuppositions and other implicitnesses are massively involved in virtually all interaction with the world, and can involve massive distortions of the individual's life in that world. That is, such presuppositions might be massively in error. This has implications both for understanding cognition and for understanding psychopathology (Bickhard, 1989). With respect to cognition, most of what an individual "believes" about the world will in fact be either presupposed or implicit, not explicit.

This is necessarily so since such implicitness can be unbounded, and in principle not capturable by explicit beliefs. No model of cognition, then, that requires that all intentionality be ontologically explicit can account for such phenomena of unbounded beliefs — such as in the frame problems (McCarthy & Hayes, 1969; Pylyshyn, 1987; Ford & Hayes, 1991; Bickhard & Terveen, 1995), or the proliferations of belief ascriptions such as "The floor will hold me up." "The chair is not a bomb." "Red trucks will hurt me if they run over me." "Green trucks will hurt me if they run over me." "If kangaroos didn't have tails, they would probably fall over, unless they got real good with crutches." and so on and so on. Presuppositions and implicitness *can* account for unbounded representationality.

Similarly, if any design approach within artificial intelligence presupposes that all valid potential intentional ascriptions must be captured in explicit representations, as most current approaches do, those approaches will be intrinsically incapable of accounting for unboundednesses of representationality.

My claim at this point is not that I have provided ontologies sufficient for all representationality — in fact, I have several more levels to go yet in this paper — but, rather, that already we are in a position to see that assumptions of singular, almost always explicit, ontologies underlying all forms of valid ascriptions of representationality will make adequate philosophy and psychology (e.g., personality theory) *and* cognitive science and artificial intelligence impossible. Already, with just five levels of representationality, there are modal and counterfactual properties of these ontologies that cannot be captured in standard approaches, yet instances of manifestations of such ontologies are arguably all construed in canonical form, in terms of beliefs and desires, thus seeming to support such assumptions of a singular common underlying ontology.

Level 6. In level 6, I introduce processes of the *apperceptive* "filling-out" of our world, which yields (in level 7) a characterization of *perception* as a specialized version. This involves, among other things, a recognition of the potentially great complexity of organization of the interactive predications about the environment.

The indications or functional interactive predications that have been illustrated to this point have all been singular and isolated. More commonly, however, we might expect the indications based on one final state to be multiple — **F87** might indicate *many* potential further procedures — and also contextually dependent on other such internal states: **F87** may indicate **P232** under some conditions, but indicate **P1033** if **F378** is also set. Furthermore, such context dependencies of functional indication could be complexly iterated and ramified throughout the indications that are functionally available for the system. **F87** may indicate the potentiality of **P232**, for example, which might, in turn, indicate the potentiality of **P92**, should **P232** be engaged in and successfully completed.

With such potential complexity of ultimate functional predications relative to the initial differentiating final states, problems of computational time and resources can become important. In particular, it may serve such a system to compute standard or default such indications, and to keep them updated, ongoingly — just in case they are needed. If they are needed, then, for some goal directed interactions, they will be already available and not require additional time and computational resources at that time.

An organization of such default indications constitutes the system's information concerning which sorts of interactions are available, and which might become available if certain other interactions were engaged in first. Elsewhere I have called such an organization a *situation image* (Bickhard, 1980b; Bickhard & Richie, 1983; Bickhard & Terveen, 1995). Creating and maintaining the information in a situation image is the functional reason for engaging in differentiating interactions, but the indications in a situation image can depend complexly not only on current and currently contextual differentiating states, but on other indications already present in the situation image as well. That is, the ongoing updating of the situation image can not only manifest contemporary context dependency, but also temporal context dependency.

The information in the situation image is updated on the basis of interaction outcomes, but goes far beyond any particularities of those outcomes. In effect, the filling out of the situation image is a filling out of information about the interactive properties of the environments, both presently available such properties, and ones available contingent on intermediate interactions. Thus, a visual scan of a glass of water indicates the potentialities for further such scans, for taking a drink, and for pouring out the water and filling the glass with diet Coke. It also indicates the potentialities for turning the glass around and scanning the printing on the back of the glass, or the colors in the glass. Similarly, a scan of a book indicates possibilities of further scans, turning the book over and ensuing scanning possibilities, opening the book and reading, shelving the book, throwing the book, and so on. The process of ongoingly updating the situation image concerning such interactive potentialities is called *apperception* (Bickhard, 1980b).

The processes that constitute apperception — the processes of constructing new indications and changing old ones — are engaged in by various apperceptive procedures. Apperceptive procedures constitute the system's knowledge of what sorts of relationships hold among various possibilities of interactions and interaction outcomes. In effect, they constitute the system's knowledge of the organization of what's possible in the world, while the situation image per se constitutes the system's knowledge of the current situation within that organization of possibilities (Bickhard, 1980b).

Default situation image will always be apperceptively constructed in such a system, but the *possible* apperceptive constructions will always exceed any *actual* situation image constructions — apperceptive indications could, in principle, always be extended further. For example, it will always be possible to trace still further such possible trajectories as emptying the water from the glass, filling the glass with Coke, offering the glass of Coke to someone, winning their undying romantic interest thereby, and so on. The constructed situation image is called the *explicit* situation image, while the potential constructions constitute the *implicit* situation image (Bickhard, 1980b).

Apperceptive processes are constructive processes, constructive of the situation image. Like learning constructions of basic system organization, there is no apriori guarantee for such constructions — they do not have foreknowledge, and are not infallible. Apperception, then, is a variation and selection constructive process.

The implicit situation image is implicit and unbounded. It is a version of the implicit manner in which potential interactive properties of an environment are implicitly defined by the procedures that would engage in those interactions. In this case, potential further apperceptive constructions are implicitly defined by the procedures that would engage in those constructions.

This unboundedness is still another instance in which assumptions of explicit representational ontologies are necessarily inadequate. Yet, once again, *instances* of both explicit and implicit situation image information are construed in canonical belief-desire terms. I *believe*, for example, that there is a back to this computer, even though I cannot see it right now; that there is a back to that book, and that that back and also the front of that book will remain even if I look away or put it in another room; that if I walk out the front door, I can see a tree; and so on. Furthermore, I believe that if I pick up this book, nothing else about the room will change; that if I move this book, its color will remain the same; and so on — classic frame problems (Bickhard & Terveen, 1995).

Note that there is, at least at a first level of analysis, a difference between presupposed representationality and situation image implicit representationality. Both can be unbounded, and, thus, can be not capturable by explicit representational ontologies; but the situation image relationships, explicit or implicit, all *involve* presuppositions concerning the conditions under which those relationships *would* hold.

Thus, the scan of the glass of water *indicates* the potentiality for taking a drink, *presupposing* that the glass isn't glued to the table, that it isn't attached to a pressure switch and a bomb, that it's really water and not acid, and so on. Part of the confusion that is invited here is due to the fact that we can construct explicit situation image representations of *some* (finite number) of our own presuppositions.

Level 6 representationality yields apperceptions, and the "beliefs" that underlie and are created by that process. It yields the complexities of context dependency and temporal state dependency, and, again, the "beliefs" that are involved and are created by those dependencies. Apperceptive potentialities are themselves implicit and involve their own implicit presuppositions, and both are intrinsically unbounded. Neither can be encompassed within an explicit finite representational ontology.

Level 7. Levels 1 through 6 capture more and more of our familiar and paradigmatic senses of representationality, but not yet enough to capture the core of much of that paradigmatic representationality. In particular, our familiar world is furnished with objects in space and time and connected by causal relations, and so on. Although I have alluded to objects in examples, the representational ontologies mentioned to this point do not account for object representation and related representation. Level 7 introduces representations of invariances such as objects; these are the representations of our more familiar world of the adult.

The implicit situation image — and, thus, the explicit situation image — would be very limited if not for the indications of, and correct presuppositions of, a great deal of continuity and unchangingness in the patterns of potential interactions. If I could not presuppose that the next room, my office, the ground outside, the streets, and on and on would remain relatively as I last directly (generally visually, in walking or driving) interacted with them, as well as similar points concerning the social relationships and institutions within which and upon which I live much of my life, then my apperceptive world would stop essentially at the indications derivable from immediately past interaction final states. There may in fact be living systems with similarly limited situation images — perhaps paramecia, maybe even frogs — but the world involves sufficient redundancy and constancy that we are not restricted in that way.

In particular, insofar as the typically encountered environments permit, it will generally be advantageous to a system to discover types of organizations of interactive potentialities that tend to remain constant, to remain unchanged or invariant *as patterns* with respect to most interactions that the system could engage in. Thus, the basic potentialities of a book remain invariant under all kinds of manipulations, movements, transportations, loanings (usually), storage, and so on. A given manipulation may change the immediately available visual scan potentialities of the book, but the original scan is still a possibility, is still available in the overall pattern of potentialities, contingent on the proper reverse manipulation that would bring the book back into the same aspect.

The invariances are not with respect to all possible interactions, however. If I burn the book, then most of its original interactive potentialities are lost. The suggestion here is that physical objects are — epistemologically, apperceptively (not metaphysically) — invariances of patterns of potential interactions under certain classes of physical interactions. Other kinds of entities will be constituted as invariances with respect to other classes of interactions, and differing kinds of entities will apperceptively manifest differing kinds of further interactive potentialities. This is a basically Piagetian account of the representation of objects (Piaget, 1954), and I would similarly suggest a roughly Piagetian account of other constituents of our familiar world, such as causal relationships, and so on.

The adaptive advantage of discovering and representing such invariances, again, is that it permits the situation image to expand without explicit bound beyond the immediately available interactive environment. Thus, it permits consideration of possibilities and contingencies that can be far removed from immediate here and now.

Note that a system that comes to be able to represent such invariances in its situation image is in effect representing temporal continuities and temporal trajectories of the environment — which, in turn, is constituting a system *memory* and *tracking* of those potentially distant potentialities. Similarly, insofar as a system came to be able to *re*construct some process of its own temporal apperceptive trajectory, it would be reconstructing its own temporal flow of activity and experience — it would be remembering its own temporal flow of activity and experience. One reason why that might offer adaptive advantage is that, because the implicit situation image cannot be exhausted by the explicit situation image, there will always be further apperceptive constructions that might be useful, or might come to be useful later, but haven't been explicitly constructed. If the system can reconstruct its past apperceptive processes, then it can carry those apperceptions out farther and in differing directions than was the case in the original flow of activity and experience.

Note that any such temporal or "event" memory will be dependent on the apperceptive procedures that participate in carrying them out — this yields a model of memory as constructive, not just a matter of retrieval. Consequently, if those apperceptive procedures have changed, then the reconstructed "memory" might similarly change (Piaget & Inhelder, 1973).

Note further that, to this point, no mention has been made of perception. In fact, there is no *input phase* of perception in this ontology (Gibson, 1966, 1967, 1979; Bickhard & Richie, 1983). All interactions change some things in the environment, and depend upon other properties in the environment. Some interactions, such as visual scans, are engaged in primarily for the differentiations they provide, and, thus, for the assistance they provide to the apperceptive process. When we find anatomically and physiologically specialized subsystems for engaging in such apperceptively specialized interactions, those interactions constitute our paradigmatic perceptual processes. But, in this ontology at least, such perceptions do not provide inputs to the rest of the system, but, rather, provide internal final state outcomes that differentiate and implicitly define, and upon which context dependent apperceptions can proceed.

Still further, some interactions that are *not* restricted to specialized subsystems, such as performing the brown ring test for iron in qualitative chemical analysis, or listening to the results of various sorts of chest thumps through a stethoscope, or lightly rubbing a finger to check for texture, or tapping a cane in various manners and patterns to determine what lies ahead, or, for that matter, moving one's head and body to obtain clarifying sweeps of view and parallax information, all nevertheless perform the same function of differentiating the environment. They are all functionally apperceptive, and, thus, perceptual, except that they surpass or transcend the specifically specialized perceptual subsystems (Bickhard & Richie, 1983). On the other hand, even the specialized systems can be used for the changes their interactions induce, as well as for their detections, such as an eye movement used as a signal. Perception, then, is not an input phase to the system, but an aspect of the environmental modulation of system activity.

Level 6, then, gives us situation images and apperception. Level 7 provides objects and other invariances that permit extensions of apperceptions beyond the immediate. It yields the emergence of two forms of memory: of environmental continuities and of system flows of activity. It captures standard paradigms of perception, but in distinctly non-standard ways.

Clearly we ascribe intentionality and representationality to manifestations of these sorts. These ontologies and their manifestations, in fact, are now beginning to capture the central forms of canonical representationality. Yet the ontologies that I have outlined that yield these manifestations are still quite different from the canonically reified ontologies of processing on encoded atoms of representation. This model, at least so far, looks exceedingly unlike anything Fodor would like — yet its instances will arguably manifest many of the forms of intentionality that Fodor wants.

I would in fact claim that these are at least much nearer the correct ontologies for the phenomena mentioned than are standard representational ontologies. But the primary point here is not so much that these are the correct ontologies as it is that we continue to encounter non-standard representational ontologies that seem to manifest classical, paradigm representational phenomena. Any system with such ontologies that we actually encountered would be described in the same canonical manner as all other representational ascriptions. Note again, however, that the unbounded capacities of the implicitnesses and presuppositions of these ontologies will unboundedly exceed the capacities of standard, finite, explicit, encoding ontologies. In general, the claim that representationality does not involve a singular common form of underlying ontology, or, at least, does not *necessarily* involve such an underlying ontological commonality, is already well made.

Level 8. Level 8 introduces the interrelated notions of off-line processing and learning. The system organization described so far is always and intrinsically in process, as is life: if life stops, it ceases. Apperception and interaction change continuously. There is no ability, in the system ontologies discussed to this point, for off-line processing. In particular, if any off-line processing were "attempted", the organizations that might be processed — the situation image — will be changing out from under it, and, further, any consequences of that off-line processing might destroy useful information in the situation image.

In level 8, I am aiming at precisely such an off-line processing ability, but there is a preliminary question to address first: why would off-line processing be useful to such an interactive system? What adaptive value would it have? After this question is answered — with a notion of why off-line processing might ever evolve — I can consider what sort of system ontology is required for off-line processing.

If a living system were optimally designed for successful interactions in its expectable environments, were optimally adapted to its niche, it would have little need for off-line processing. In fact, since off-line processing requires resources for stepping off-line in some way — both time and computational resources — optimal adaptedness precludes off-line processing. It only gets in the way and wastes resources.

If a system were optimally designed for successful interactions in environments characterized by change and novelty, however, the design criteria are different. That is, optimal design for adapt***ed***ness is not the same as optimal design for adapt***ive***ness — the ability to learn. In fact, I argue that optimal design for adaptiveness, for learning, does involve off-line processing.

Level 8, then, involves systems that learn. I am not going to focus here on the learning per se, however, but rather on some design principles that affect the interactive and representational organization of a system that is capable of learning. In other words, learning is not only an *addition to* a simple interactive system, it also changes the nature of optimal organization within that interactive system.

Note first that, in the ontologies discussed, representationality is emergent in system organization, emergent in the ability of various system organizations to take into account differentiations of the environments in further interactions with those environments. That is, representation is not constituted in encoded correspondences with items and properties in the environment; representation is not constituted in any kind of structural relationships with the environment, but, rather, in functional interactive relationships. With respect to learning, a major consequence is that there is no temptation to hold a version of a passive imprinting into the system as a model of learning. Synchronic transduction and diachronic induction are simply passive imprinting or scratchings into a wax slate, that the system is then somehow supposed to understand. The homunculus required to interpret any such transductions or inductions, of course, simply recapitulates the basic problem of representationality that was supposedly being addressed.

For the interactive system ontologies, however, there is no temptation to suppose that the environment could possibly imprint or scratch-over-time successful system organizations. New system organizations can only occur via internal system constructions. Learning in these ontologies, then, must be some form of constructivism.

Furthermore, although there will in general be many cases in which such internal construction can make use of prior knowledge concerning what sorts of things might work, ultimately it is not possible to know ahead of time what constructions will work. Prescience or foreknowledge eliminates the need for learning, and learning constructions that are prescient concerning what constructions to engage in is not learning at all — the knowledge is already available. The system cannot advance beyond its current knowledge with knowledge of where to go. The constructions in these ontologies, then, must be trials that might or might not work. Learning must ultimately be some form of variation and selection constructivism, some form of evolutionary epistemology (Campbell, 1974, 1990; Hahlweg & Hooker, 1989; Hooker, 1995; Popper, 1965, 1972; Radnitzky & Bartley, 1987).

It is the construction process that forces new design principles for optimality in learning systems. In particular, in a system that is optimal for interaction, with no learning, there might conceivably be two or more copies of some strategy of interaction located in differing parts of the overall system. A programming esthetic would want to pull that strategic organization out into a subroutine, or some equivalent — a single copy of the strategy that could in principle be used from multiple locations in the rest of the system. For interaction per se, however, that simply incurs additional resource costs — optimality for adaptedness involves in-line copies of everything, even if that involves some degree of duplication (it is conceivable that some interactive organization would be used from so many different other locations in the system that it would be optimal to lift it into a subroutine single copy, in spite of the computational costs, in order to save the memory, or nervous system, resources — I would not expect that to be common).

In a learning system, however, if all learning constructions were of directly in-line system organizations, then new copies of particular organizations would have to be reconstructed from scratch at each point in which they were needed. “Everything in-line” precludes versatile re-use of previous constructions in new constructive attempts. “Everything in-line” makes learning extremely inefficient: it is not optimal for a learning system, however much it may be optimal for a non-learning interactive system.

"Everything in-line" requires that the wheel be reinvented every single time it is needed in a new part of the system.

Optimality for learning, thus for adapt*ive*ness, requires that previous constructions be re-usable for new construction attempts; it requires that the resources available for learning be themselves learnable — constructable and re-usable — a recursiveness of construction. It requires that there be at least this minimal form of *learning to learn* (I dub this a minimal form because, although it involves constructing *resources* that are usable for new constructions, it does *not* involve improvements in the constructive processes themselves: see Bickhard, 1992b).

But, in order to make such use of already constructed organizations in new constructions requires that those organizations somehow be available off-line, at least initially, in some sort of off-line resource cache or domain that can be made use of in such new construction attempts. Optimal learning, then, requires off-line system organizations — general subroutines or objects or agents of some sort. Off-line possibilities then, improve learning — learning as itself a quasi-evolutionary variation and selection constructive process — and that is why they might be expected to evolve. Note that subroutines, objects, and so on in programming have not been developed for their consequences for the *programs*, but, rather, for their improvements in the *programming*; the history of programming and of programming languages has recapitulated the selection pressures and their consequences that I have been discussing here.

The next question, then, is how such off-line processing could work within the system ontologies presented. To try out an off-line subroutine resource in interaction requires that that off-line system organization be somehow given access to the relevant information in the situation image. A programming perspective would suggest either copying that information for the subroutine, or providing pointers to that information. Neither strategy will work in the general case.

Providing pointers to the potentially relevant parts of the situation image for use by an off-line system might work if the overall system were either not parallel or was not continuously in interaction. But, for living systems at least, and for any systems that are parallel and/or are continuously in interaction, pointers fail because there is no apriori guarantee that the situation image will stand still while the off-line system does its processing. The situation image is the primary resource for ongoing interaction for the system, and, unless that interacting can somehow be halted, the situation image will be continuously changing in updates from and for those interactions. Pointers, then, won't do.

Copying information, copying organizations of indicators, from the situation image and providing such a copy to the off-line system might seem to solve this problem. In a sense, it does, but it encounters its own problem — a problem that would hold even for non-parallel systems whose interacting could be halted. This problem is that there is no apriori guarantee that the information will be in the form that the off-line system needs. By its nature, the off-line system and the domain of interaction in which that off-line system (or some construction of off-line resources) is being tried out were not designed for each other. In fact, in the local opportunism typical of evolutionary and quasi-evolutionary constructive processes, it can be expected in general that differing parts of the overall system will be potentially quite **ill**-designed for each other. Copies, then, won't work because they can in general be expected to be of the wrong form for the off-line resource to use.

What is required in the general case, then, in order for off-line resources to be generally useful for further constructions, is some set of procedures that can construct *translations* of situation image organizations into forms useful for the off-line resources. That is, optimality in a learning system involves off-line construction resources, which, in turn, involves procedures for constructing, not copies, but more general stand-ins for situation image organizations. These stand-ins could be provided to the off-line resources for computation, and the results of those computations then translated back into whatever form was needed within the situation image. Optimality in learning, then, requires off-line resources, which require informational stand-in constructors. Note that stand-in construct***ion*** is something that the overall system can potentially get better and better at — in learning and development — by constructing more powerful and more specialized stand-in construct***ors***. If the off-line constructions prove useful and are recurrent, it might prove additionally useful to move those off-line constructions into integrated on-line copies (Bickhard & Richie, 1983).

Note also that many such translations will always be logically possible, and that there is no apriori information about which of those possible translations will be useful (though there may be previously learned heuristic such information). Stand-in construction, then, will also be a variation and selection process.

My current focus, however, is on the stand-in-constructors per se and on what they construct. Informational stand-in constructors, I claim, are internal, functional encoding constructors. They are translators of the information in the situation image into different forms appropriate to the off-line resources. These are encodings in generically the same sense in which paradigmatic encodings are, such as Morse code. In Morse code, "**. . .**" stands-in for " **S**"; it changes the form of "**S**" in useful ways, since "**. . .**" can be sent over telegraph wires, while "**S**" cannot. Similarly, bit codes do the same in computers, permitting many things that would be impossible otherwise.

As mentioned earlier, I argue that such stand-in encodings are the only form of epistemic encodings possible, and that assuming encodings can be anything *other* than stand-ins is a source of enormous confusion (Bickhard, 1992a, 1993; Bickhard & Terveen, 1995). In fact, canonical belief-desire psychology, Fodor, artificial intelligence, connectionism, cognitive science, psychology, and so on, all *do* presuppose that encodings can be foundational *providers* of representational information, not just stand-ins for representational content that is already available — and they *are* afflicted with consequent confusions (Bickhard, 1992a, 1993; Bickhard & Terveen, 1995; Winograd & Flores, 1986; e.g., Laird, Newell, Rosenbloom, 1987; Newell, 1980a, 1980b; McClelland & Rumelhart, 1986; Rumelhart & McClelland, 1986; Rumelhart, 1989; Waltz & Feldman, 1988). Note that the representational content for the encodings defined here is provided by the situation image, which, in turn, has representational content by virtue of the functional predications of implicitly defined interactive properties that make up the situation image.

With the recognition that adaptedness for learning will tend to select for encodings, whether in biological evolution or artifactual construction, we encounter an ontology that looks startlingly familiar. It looks, at least superficially, like classical computer models of processing of encoded representation. Translators construct representational stand-ins, encodings, which are then processed into new encodings — inference: whether demonstrative, nondemonstrative, monotonic, non-monotonic, and so on — which are then translated back into the situation image for trial in interaction. Level 8, then, yields, or seems to yield, familiar encodings and the processing of encodings. Only the initial and final steps of translation seem odd from the standard perspective.

Those initial and final steps, however, are of critical importance. Level 8 ontologies manifest learning, and manifest the processing of encoded information, but the requirement that those encodings be stand-ins for situation image representations entails that those encodings will not be standard, and will not have standard properties. Standard information processing encodings, in fact, are presumed to carry representational content of what they represent, and they are generally presumed to represent by virtue of what they correspond to (Fodor, 1987, 1990; Hanson, 1990; Newell, 1980a, 1980b; Bickhard, 1992a, 1993; Bickhard & Terveen, 1995). The encodings of level 8 system ontologies are derivative from situation image organizations, however, and will inherit whatever properties of representationality that are inherent in that interactive representational base. This involves, among other things, *implicit* definition of environmental categories, *implicit* definition of interactive properties, *functional* rather than syntactic predication of those interactive properties, and so on.

Furthermore, the *internal* processes of translating from the situation image to some organization of internal stand-ins, and the converse of translating back into the situation image, are strictly *functional* processes; *external* encodings, in contrast, involve intentional interpretations, not just functional translations — we must *understand* the Morse code stand-in relationships in order to make use of "**. . .**" for " **S**". If the *internal* translations required the same sort of intentional interpretation as do *external* encodings, then their existence and use would require the same explanation of intentionality that was being attempted in the first place — the first step would be taken on an infinite regress. This regress that is involved in presupposing that external representationality could serve as a model for internal representationality — the necessity for internal interpretive homunculi — is still another aspect of the flawed assumptions in standard approaches (Bickhard, 1993; Bickhard & Terveen, 1995).

The difference that I wish to focus on here is that interactive representation is fundamentally and necessarily a matter of interactive differentiation and implicit definition, rather than of an inputting of transduced particulars into an encoding processing system. Interactive differentiation and implicit definition are always of interactive *categories*, never directly of particularities. This holds even for objects and individuals: the claim or assumption that there is only one member of some interactive category, that the extension of an interactive category is a unit set (a set with only one element in it), is in addition to and beyond the definition of that category per se. Because some forms of reference involve such presuppositions of differentiation to individual uniqueness (not all forms of reference do: "The whale is a mammal."), this point will affect reference as well.

The basic task of learning in this ontology is not to produce generalization, as it is with standard empiricist encodingisms, but the converse — differentiation. The dominant task is to construct ever finer and more finely relevant differentiations of environments, with differentiation sufficient to yield unit set extensions, differentiations of individual objects or entities or properties, constituting (sometimes) important criteria to hold and accomplishments to constructively aim for. Implicitly defined categories are already, and intrinsically, general (though, of course, they may not be the *right* generalities for some task, in which case the construction of more or different general differentiations can also be attempted — but even this will not be *generalization*, from *particulars*, as understood within empiricism; Loux, 1970; Bickhard & Terveen, 1995). It is not encodings of individual objects that are input to the interactive system, they are not given as the ground of representationality — in fact, *nothing* that is *epistemic* is *input* to an interactive system at all. Differentiations sufficient to such unit-set-ed-ness must be *attempted*, not given or presupposed, and system organizations sufficient to such differentiations must be constructed.

Furthermore, any such differentiation to unit set-ed-ness, of single individuals, is a *defeasible* accomplishment (a critical principle; Bickhard, 1991a, 1992a, forthcoming, in preparation-a). It is always possible to discover that we have been mistaken, that more than one instance of a differentiation either does or can exist. Differentiation of individuals is not guaranteed, precisely because *differentiation* is the fundamental epistemic function, not inputs of encodings that are already of particularities. Individuals, and their differentiating specification, must be constructed.

These representational ontologies are deeply non-empiricist in their representationality. They differ from an empiricist approach in two fundamental ways. First, representationality is *constructed* in this ontology; constructed as properties of interactive system organization. In particular, it is *emergently* constructed as functional organization, and does not require constructive raw materials that are themselves already representational. Interactive representation does not have to come from *anywhere*, and certainly not in through the senses or up from the genes; with respect to representationality per se, interactive representation is constructed *de novo*. Interactive representationality does not require representation in order to construct representation (Bickhard, 1991b, 1992a, 1993; Campbell & Bickhard, 1986).

A second point of difference with empiricist epistemologies concerns the lack of particularities. Individuals having to be defeasibly constructed and differentiated, rather than being taken as the epistemic ground, is fundamentally at odds with the empiricist tradition, in which standard artificial intelligence and cognitive science, as well as related domains in philosophy and biology, are deeply embedded.

If these interactive ontologies are necessary, relative to standard encoding computational models, then those standard computational models involve many basic errors in assumptions and presuppositions. I argue that interactive ontologies *are* necessary, and that standard architectural assumptions about representational ontologies are at root logically incoherent (Bickhard, 1992a, 1993; Bickhard & Terveen, 1995). My focus here, however, is simply that these are plausible architectures, plausible system ontologies, that would manifest various forms, even seemingly canonical forms, of representationality. If these non-standard ontologies are in this sense at least *sufficient* to representationality, setting aside issues of whether they are necessary, then standard approaches are at a minimum incomplete, and in error insofar as they claim or assume that such standard encoding architectures are the only correct approach, or are "all we've got." (Fodor, 1975, 1981, 1987, 1990). I argue, as noted, that they are not just incomplete, but that they are incoherent (Bickhard, 1992a, 1993; Bickhard & Terveen, 1995). Still further, these interactive ontologies have already been shown to be more powerful than explicit encoding ontologies, for example, by virtue of their implicitness and consequent potential unboundedness. There are many additional arguments for their superior power (Bickhard & Terveen, 1995).

In level 8, then, we find the emergence of something that superficially looks like standard computational models operating on encoded symbols. This will be manifested in such representational and developmental phenomena as the availability of progressively broader and more powerful general strategies, the ability to engage in off-line processing preparatory to further interaction, and so on. But a closer look shows that what is being processed in level 8 ontologies is representationally fundamentally different than is assumed in standard computational approaches.

Level 9. In level 9, I wish to discuss language, though mostly with respect to some negative points. The negativity derives from the fact that standard conceptions of the ontology of language cannot be constructed within the framework that has been developed, though a system ontology whose manifestations would look exactly like language can be.

In standard views, utterances are construed as encodings of mental contents which are transmitted into another mind where they are decoded into the receiver's mental contents. The only sort of encodings that the ontology presented can support, however, is stand-in encodings, and stand-in encodings can only stand-in for other representations that the system already knows about, already has functional access to. In terms of stand-in encodings, an audience to an utterance, in order to decode that utterance, would have to already know what the words in that utterance stood-in for in the speaker's mind. Stand-ins can't provide new representation, they can only recode representation that is already present.

It might seem that such stand-in encoding relationships from utterances to mental contents could be learned as part of language learning, but, skipping over the arguments, it turns out that any such learning requires some foundation upon which it could build, and such a foundation simply recreates the problem: in order to learn what the stand-in is a stand-in for, the child would have to already know what is being stood-in for (Bickhard, 1980b, 1987). Stand-ins require representation in order to get representation; they are not creators of representation.

In sum, in the ontology that I have been presenting, language cannot be encodings of mental contents because no ontology for such new-representation-creating encodings has been presented. Again, other arguments yield the conclusion that no such representation-creating encodings are possible (Bickhard, 1992a, 1993).

A second reason why the interactive ontology cannot support the standard model of language is that the only sort of relationship between a system and its environment that this ontology *can* support is that of some form of interaction. In standard views, most activity of an animal in its environment, including that of human beings, is interactive, but language is taken to be a rather mysterious exception. Instead of an *interaction*, an utterance is presumed to be a *transmission* — of an encoding. This notion of a transmission is not renderable in interactive terms. To mention just one problem, transmissions have contents, representational contents, and interactions do not.

Utterances in the interactive ontology must be some sort of interaction — but what sort? How are language interactions different from other kinds of interactions? Skipping the arguments again, I propose that language interactions are distinguished by the objects of the interactions, by what they interact with. Specifically, I propose that utterances interact with social realities — they operate on and transform social realities of conversational flow, of discourse flow, of conventions, of institutions and activities within the framework of institutions, of relationships, and so on. Social realities are understood to be constituted as certain forms of commonality of understanding of the situation by the participants in that situation. Utterances, then, operate on and transform — just like most other interactions — but what they operate on and transform are social realities constituted as commonalities of understanding: situation conventions is the term I use for such commonalities (Bickhard, 1980b, 1987, 1992a).

Systems can constitute parts of each others environments, and, when they do, the situation image of that environment takes on a special problematic character for each of the systems involved. In particular, characterizing the situation involves characterizing the other systems, but the interactive potentialities of the other systems depend in turn on their characterizations of the situation too, including their characterizations of the first system (and each other). Since each system does have an interest in constructing a situation image of the situation, including of the other systems, but no system can accomplish that without arriving at some commonalities among the systems' situation images — commonalities that can constitute a halt to the regress of characterizing the other system's characterization of my characterization of his characterization of ... and so on. This common interest in arriving at a necessarily joint or common characterization — a mutually consistent characterization — constitutes a coordination problem (Schelling, 1963), and a solution constitutes a convention (Lewis, 1969), a situation convention (Bickhard, 1980b).

This much follows simply from the nature of such systems as being epistemic agents. Characterizing the situation in the situation image, then, requires, among other things, characterization of the representationalities of the other systems, but those will involve characterization of the representationality of the first system. This yields reflexivity and the coordination problem. Solutions will generally be in all system's interests, though how to arrive at such solutions is not always easy.

Such solutions, situation conventions, furthermore, will not remain static. Systems will be themselves be changing and will have interests, potentially, in creating new situation conventions among these or new such agents. The interactions of systems can be apperceived by other systems, and this provides a potential manner in which systems can interact with, transform, operate on those situation conventions. Language, in this view, is itself an institutionalized convention for the manipulation and transformation of situation conventions; utterances are interactions with the environment that constitute interactions with situation conventions by virtue of the joint apperceptions of those utterances (Bickhard, 1980b).

Utterances of this ontology will change the representationality of each of its audience's and its utterer's situation images. Utterances will alter the representations in the situation images, eliminating some and creating others. In this way, utterances can have the result of creating new representation in an audience — a similar result as for the standard transmission models, but accomplished in a very different way. There are strong intrinsic constraints on how this could be accomplished that are manifested as syntactic regularities among utterances — again, a similar result as for standard language models, but derived in a very different way: a much more natural way (Bickhard & Campbell, 1992; Bickhard, 1995a).

Furthermore, utterances will accomplish these alterations in an intrinsically context dependent manner: an utterance will be apperceived within the context of already existing situation images, within the context of already existing situation conventions, and, thus, those contexts will be determinative along with the utterance itself of the apperceptive consequences. Utterances are operators on situation conventions, and the results of operators depend as much on what is being operated upon as on the operator per se. Such context sensitivity is in fact ubiquitous in language, but accounting for it is ad hoc within standard utterances-as-encodings models (Kaplan, 1979a, 1979b, 1989; Richard, 1983; Fodor, 1987, 1990; Loewer & Rey, 1991). Context sensitivity is *intrinsic* in the interactive ontology (Bickhard, 1980b; Bickhard & Campbell, 1992).

Utterances cannot be simply *input* as operators: this conception assumes that utterances are to be decoded. Utterances must be apperceived, just like all other influences from the environment, and apperception is intrinsically constructive and fallible. The understanding of utterances, then, is a variation and selection, trial and error, constructive process. It is only in simple cases of language understanding that it might appear otherwise: in circumstances of difficult language understanding — such as children's language learning, interpreting historical texts, or in psychotherapy — the difficulties and the fallibilism, the trial and error, become clear.

Most importantly for my current purposes, if we were to observe systems that could and did interact with each other's understandings of the joint situation in this manner, then, so long as those interactions had a productive form, we would strongly tend to ascribe the intentionality of language to them. In the interactive view, utterances are not themselves directly representational; instead, they *operate on* representations. But they do thus create new representations, and it is quite easy to mistake the context sensitive transformations of representations into new representations as the transmission of those new representations (Bickhard, 1980b).

In level 9, then, an ontology with manifestations to which we would ascribe language emerges, and emerges naturally in terms of an intrinsic coordination problem created by the mutual considerations of epistemic agents. As is by now familiar, the ontology is quite different than in standard views, and so are some of the more subtle properties of the utterance manifestations, but the general phenomena of the interactions of this ontology capture what language looks like and what language is used for (for deeper explications, see Bickhard, 1980b, 1987, 1992a; Bickhard & Campbell, 1992; Campbell & Bickhard, 1992a).

Level 10. Level 10 is concerned with conscious representationality and conscious goals. Once again, I will sketch an ontology rather than develop it in any detail. My purpose here is to outline a general architecture rather than the details of that architecture. I will focus on consciousness in the sense of reflexive consciousness, not just in the sense of "not asleep", or "not in coma".

Fundamentally, I propose that reflexive consciousness is reflexive awareness — awareness of processes and properties of awareness. Awareness, in turn, is ongoing interaction with an environment by an ontology of perceptual and motor systems engaged with and by the apperceptive processes of a situation image. That is, I propose that awareness — experiencing — is the activity of the basic ontology already outlined in the first nine levels of representationality. The objects of reflexive awareness, consciousness, are the organizations of guises or presentations of the sort mentioned in the level 8 discussion, and those organizations are the organizations of the activities that constitute first level awareness.

Still more specifically, consciousness, in this view, is the interaction by a second level interactive knowing system with the first level interactive knowing system *that has been sketched up to this point in the first nine levels of representationality.* In the same sense in which an interactive system can manifest representationality with respect to its environment, a second level knowing system taking the first level system organizations and processes as its environment can manifest (constitute) representationality with respect to that first level system — can manifest reflexive representationality.

With such an architecture, the overall system can manifest many new properties. It can hold particular first level representations in interaction (rehearsal); it can make use of first level organization as an image of how the environment would react to various system activities, thus allowing planning ahead (Bickhard, 1978); it can differentiate between the interactive representationality evoked by the apperceptions of an utterance and the situation image organizations that constitute that representationality — that is, it can distinguish between an utterance and the utterance's meaning — and so on. Whether or not this architecture is accepted as constituting reflexive consciousness in its activities, it is clear that interesting and adaptive new possibilities emerge, and that we standardly tend to ascribe them to consciousness.

Furthermore, this architecture yields a potentiality for ascending even higher up a hierarchy of knowledge about knowledge about knowledge, and so on. Each level of interactive knowing will instantiate properties that might be useful to represent, to know, and that can be interactively known from a next higher level. In addition, each higher level will instantiate properties of system and process that are more abstract than those of the levels below. The first level system can represent its environments, while higher levels could represent, for example, the ordinality of iterations of a lower level procedure. In other words, higher knowing levels offer knowledge of progressively higher abstractions, such as mathematics (Campbell & Bickhard, 1986). (Note that these levels of *knowing* extend in a dimension orthogonal to the levels of *representationality*, and include the first nine levels of representationality at the base of the knowing levels.)

I argue that the first instance of ascent up these knowing levels, the possibility of interactive functioning at a second level of knowing or awareness, must emerge as an architecturally explicit second interactive level of system — an evolutionary emergence — but that higher levels can be ascended in a purely functional manner given the initial two levels (Bickhard, 1973, 1980a; Campbell & Bickhard, 1986).

Since all learning and development in this ontology must ultimately rest on variation and selection constructivism, there is no possibility for a system to develop in such a manner that it skips some level of knowing. To skip a level would be to have an interactive system with nothing to interact with at the level below it. System development, then — child development, for example — must ascend these levels, if at all, in strict sequence. These levels of knowing generate an immutable stage sequence of possible development (Campbell & Bickhard, 1986).

Since the initial ascent to the second knowing level is architectural, it should be expected that it would occur roughly at a given age, likely by virtue of neural maturation of the second level system, across most domains of representation. But ascent to further levels will be purely functional; thus, neither constancy of age of transition to higher levels across children nor across domains within a particular child should be expected on the basis of this ontology. Both of these properties seem to be borne out in the developmental data — the initial domain general age of transition (age 4) and the general lack of synchrony in further stage ascents (Bickhard, 1973, 1980a, 1992c; Campbell & Bickhard, 1986).

This ontology, then, not only captures manifestations of consciousness, it also yields predictions concerning development that seem to be supported. It is the only developmental model that makes those particular predictions, and they were first derived in 1973, long before most of the current supporting data was collected (Bickhard, 1973, 1978, 1980a, 1992c).

Reflexive consciousness in this model is not reflexive passive perception. It is not a meta-receipt of encoding inputs of other mental contents or objects — it *cannot* be. It is instead a meta-knowing interaction, a meta-representationality. For at least the tenth time, an unfamiliar ontology appears here to be capable of manifesting familiar forms of representationality, but in nonstandard ways, and with nonstandard specifics of manifestation.

I have a number of times mentioned and outlined arguments with the conclusion that standard assumptions about representational ontologies are logically incoherent, and, thus, cannot be correct. I have at most adumbrated those arguments — they have already been more fully developed in several places elsewhere — but have focused instead on the delineation of system ontologies that plausibly manifest various forms of what would appear to be representationality. This focus on alternative ontologies complements the critique of standard approaches: insofar as these system ontologies would manifest something that looks very much like representationality, and various levels of representationality at that, then proponents of standard views must either defeat the claims of ascription of representationality to the manifestations of such ontologies, or give up claims of constituting "the only game in town" (Fodor, 1987, 1990; Newell, 1980a, 1980b). Still further, they must give up standard approaches entirely or else not only defeat the arguments of incoherence, but show how to capture in a non-ad hoc manner the context sensitivities, the unbounded implicitnesses and presuppositions, the developmental stage constraints, and so on that are intrinsic to the interactive ontologies.

The focus in this model on activity instead of on empiricist inputs (or "up from the genes" "rationalism") as the framework for representationality is shared with several contemporary approaches to intentionality. Implicitly defining organizations of potential interactions are like *presentations* of objects and individuals, or their *guises*, not the individuals themselves; the processes of interactive differentiation and indication of potential interactions are like the *experiences* or *meanings* of those organizations of potentialities in the environment. In these respects, the model is akin to the Kantian systems of Meinong or Twardowski or Castañeda (Castañeda, 1989; Santambrogio, 1990; Tomberlin, 1986), or to the interactive pragmatic and process phenomenology of pragmatism (Bourgeois & Rosenthal, 1983; Houser & Kloesel, 1992; Murphy, 1990; Rosenthal, 1983, 1986, 1987, 1992; Thayer, 1973). Furthermore, the focus on engaged interactive process shares themes with existential phenomenology: The implicit representationality of interaction systems is akin to "skill intentionality" (Dreyfus, 1967, 1982, 1991a; Dreyfus & Dreyfus, 1986, 1987; Bickhard & Terveen, 1995). The constructive interpretative model of language is akin to the hermeneutic circle (Dreyfus, 1991a, 1991b; Gadamer, 1975, 1976; Guignon, 1983; Heidegger, 1962; Okrent, 1988). These similarities are as much due to the shared anti-empiricism, and the shared focus on process and action, as to detailed commonalities. In other respects, not developed here, the model is quite different from these traditions. The existence of such kinships, however, renders the typical "only game in town" defense of standard "representation-as-correspondence" approaches even more vacuous.

Further Properties of the Model

Intrinsically Situated and Context Dependent

Many properties of these interactive ontologies of various forms and levels of representationality have been mentioned, such as their implicitness, unboundedness, Kantianism, pragmatism, and so on. There is one class of properties, however, that I would like to focus on a little more explicitly.

In particular, the interactive model offers support to a perspective on cognition called "situated cognition" (Agre, 1988; Chapman & Agre, 1986; Clancey, 1989, 1991; Maes, 1990). It offers that support by virtue of the fact that the interactive ontologies intrinsically manifest and force a situatedness of representationality. Thus interactivism both offers an ontology for the realization of situatedness and a convergent set of considerations that yield situated conclusions.

In a way, the point is obvious: interactive representational ontologies require intrinsically open interacting systems. Such systems represent characteristics of their environment by virtue of their interactions and potential interactions with those environments. The environmental differentiations and implicit definitions of interactivism are intrinsically relative to the system and its differentiating interactive activities — they are intrinsically, necessarily, indexical. The interactive properties that are functionally predicated in interactive representation are intrinsically, necessarily, functional — properties of the interactive functional relationships between the system and its environment, properties of the situated relationships between the system and its environment. Interactive representation, then, is intrinsically, necessarily, deictic in the situated sense outlined by Agre (1988).

Interactive differentiations are intrinsically context dependent in the sense that just what a differentiating interaction will in fact differentiate will depend on the environment in which it is engaged. Such differentiations will necessarily *tend* to be of situations appropriate to the predicated further interactive possibilities in usual environments in order for those differentiations and their predications to survive selections, either biological selections in evolution or constructive selections in learning. But in no case is there any apriori guarantee that the functional predication is correct under all situational circumstances.

Interactive representation is an emergent of action, interaction, and the potentialities for interaction; it is not a product of any sort of processing of inputs. Passive systems that only receive inputs *cannot* be representational in the interactive sense. No "processing" of inputs can yield interactive representationality, including any connectionist processing of inputs: representation is an emergent property of situated *interacting* systems, not passive systems. Neither computers nor connectionist networks, then, can be interactively representational; robots, however, could be (Beer, 1990; Bickhard, 1982; Bickhard & Terveen, 1995; Brooks, 1991a, 1991b; Cherian & Troxell, 1995, in press; Hooker, in preparation; Maes, 1990; Malcolm, Smithers, & Hallam, 1989; Smithers, 1992).

Interactive representation is an emergent of *potentialities* for interaction. It is *not* most fundamentally of or about *actualities* except in the sense of "actual" interactive potentialities. This intrinsic involvement of modality in all representation is characteristic of all recognitions that representation emerges from action, not from inputs, and interactivism shares in it. Representation as emergent from action rather than from the processing of inputs *must* be modal because it is potential actions, the potentialities for action and interaction, that are at issue, not past actions and not past inputs (Bickhard, 1988; Bickhard & Campbell, 1989; Bickhard & Terveen, 1995). Standard conceptions of representation focus on supposed encodings of actualities, derived from inputs corresponding to those actualities, with either no involvement of modality, or with an ad hoc addition of modal operators on top of a basically non-modal representational system.

Folk Psychology

This interactive framework, especially when supplemented by the arguments for the incoherence of standard encodingist approaches (Bickhard, 1991b, 1991c, 1992a, 1993), gives rise to questions about standard belief-desire folk psychology. If folk psychology isn't capturing the correct underlying ontologies, what status does it have? Churchland (1981, 1984) construes folk psychology as a theory about mental ontology, and as an egregiously false such theory. Freud and Fodor take it to be a correct theory. I suggest as an alternative that folk psychology is primarily concerned with the regularities of and with regulating the interactions of people with their environments, most especially with their social environments. Any such functionality of folk psychology will impose constraints on what sorts of underlying ontologies could be consistent with such functionalities, but it does not require that folk psychology be any sort of direct theory about those underlying system ontologies (Clark, 1989).

In particular, folk psychology is concerned with the accomplishments, and the claimed and intended accomplishments, of the interactions of people more than with the processes by which those accomplishments were or were not attained. Accomplishing truth or reference, for example, is not something that can be assured to any procedure for accomplishing them, yet notions of knowledge presuppose the accomplishment of truth and notions of naming and "talking about" assume the accomplishment of unit set differentiation. All such cases are defeasible. They are presumptions about the relationships between the organisms and their worlds, not about the interiors of the organisms per se. They are accomplishments that we ourselves and other people alike are vitally interested in, and, thus, they are an understandable focus of folk psychological characterizations and interchanges. Such accomplishments constitute an important part of the exchanges between and among people — exchanges of knowledge, or at least ostensible knowledge, for example.

Claims to such accomplishment, then, even though intrinsically defeasible, intrinsically incur obligations and responsibilities concerning those claims, *and* concerning the unboundedly implicit presumptive conditions of those claims: other people depend on those ostensible accomplishments having been actually accomplished and on those conditions in fact existing — on ostensible knowledge being true and ostensible reference succeeding. This realm of the normativity of interactions among people emerges necessarily in part because of the exchange-importance of such defeasible accomplishments.

There is, of course, another aspect to the use of folk psychology — its uses in predicting and understanding others' actions. This is more amenable to interpretation as a theory. But, at the folk psychology level per se, this use is primarily instrumental: it characterizes dispositions, and explains behavior in terms of such dispositions. To characterize the multifarious forms of presupposition, implicitness, predication, and so on, and the dispositions to which they give rise, in folk psychology terms is at best a partial characterization — one that collapses most of the levels of representationality into one canonical form and elides the implicit character of those levels — but it can nevertheless be quite powerful, and it generally suffices for instrumental purposes. When such uses are pressed into theory, however, when the notions of folk psychology are reified into ontological mental elements, we get such aberrations as the encoding incoherencies, the Dynamic Unconscious, and the frame problems.

In this view then, folk psychology is not a theory about anything (which does not preclude its being *treated* as such, by Churchland or Fodor or Freud, for example). It is a framework within which claims and exchanges of various forms of representational accomplishments and behavioral dispositions can be made and discussed. In particular, folk psychology does not provide, and should not be taken as ostensibly providing, a theory concerning the ontology of representationality.

I have argued, instead, that there are many forms of representationality and many forms of underlying ontologies for those representationalities. Interactive functional presuppositions, implicit definitions of environmental categories, implicit definitions of interactive properties, implicit definitions of potential apperceptions, and so on, all constitute forms of representationality. They do not all have the same form, and do not all involve the same underlying ontologies — and none of them is identical to, or even consistent with, the common reification of belief-desire folk psychology into propositional attitude ontologies.

Conclusion

Representationality is not a singular natural kind. It does, however, have something like a family-resemblance underlying coherence: all of these types and forms and levels of representationality are properties or aspects or organizations or emergents of situated interactive goal-directed systems. None of these forms of representationality are anything like a language of thought or a data structure or a connectionist attractor: none of them involve the notion of representation-as-correspondence-to-what-is-represented that is common to ideas of language of thought or data structures or information to be processed or connectionist "representations". Instead, interactiveness is the common thread.

References

Agre, P. E. (1988). The Dynamic Structure of Everyday Life. Ph.D. Dissertation, MIT AI Laboratory.

Bach, K., Harnish, R. M. (1979). *Linguistic Communication and Speech Acts*. MIT.

Beer, R. D. (1990). *Intelligence as Adaptive Behavior*. Academic.

Bickhard, M. H. (1973). A Model of Developmental and Psychological Processes. Ph. D. Dissertation, University of Chicago.

Bickhard, M. H. (1978). The Nature of Developmental Stages. *Human Development*, *21*, 217-233.

Bickhard, M. H. (1980a). A Model of Developmental and Psychological Processes. *Genetic Psychology Monographs, 102*, 61-116.

Bickhard, M. H. (1980b). *Cognition, Convention, and Communication.* New York: Praeger.

Bickhard, M. H. (1982). Automata Theory, Artificial Intelligence, and Genetic Epistemology. *Revue Internationale de Philosophie, 36,* 549-566.

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. (1988). The Necessity of Possibility and Necessity: Review of Piaget's Possibility and Necessity. *Harvard Educational Review, 58*, No. 4, 502-507.

Bickhard, M. H. (1989). The Nature of Psychopathology. In Lynn Simek-Downing (Ed.) *International Psychotherapy*: *Theories, Research, and Cross-Cultural Implications*. Westport, CT: Praeger Press.

Bickhard, M. H. (1991a). A Pre-Logical Model of Rationality. In Les Steffe (Ed.) *Epistemological Foundations of Mathematical Experience.* New York: Springer-Verlag, 68-77.

Bickhard, M. H. (1991b). Homuncular Innatism is Incoherent: A reply to Jackendoff. *The Genetic Epistemologist*, *19*(3), p. 5.

Bickhard, M. H. (1991c). The Import of Fodor's Anti-Constructivist Argument. In Les Steffe (Ed.) *Epistemological Foundations of Mathematical Experience.* New York: Springer-Verlag, 14-25.

Bickhard, M. H. (1991d). Cognitive Representation in the Brain. In *Encyclopedia of Human Biology*. *Vol. 2*. Academic Press, 547-558.

Bickhard, M. H. (1992a). How Does the Environment Affect the Person? In L. T. Winegar, J. Valsiner (Eds.) *Children's Development within Social Contexts: Metatheory and Theory.* Erlbaum, 63-92.

Bickhard, M. H. (1992b). Scaffolding and Self Scaffolding: Central Aspects of Development. In L. T. Winegar, J. Valsiner (Eds.) *Children's Development within Social Contexts: Research and Methodology.* Erlbaum, 33-52.

Bickhard, M. H. (1992c). Commentary on the Age 4 Transition. *Human Development*, 182-192.

Bickhard, M. H. (1993). Representational Content in Humans and Machines. *Journal of Experimental and Theoretical Artificial Intelligence*, *5*, 285-333.

Bickhard, M. H. (1995a). Intrinsic Constraints on Language: Grammar and Hermeneutics. *Journal of Pragmatics, 23*, 541-554.

Bickhard, M. H. (1995b). Who Interprets the Isomorphisms? *New Ideas in Psychology*, *13*(2), 135-137.

Bickhard, M. H. (forthcoming). Critical Principles: On the Negative Side of Rationality. In Herfel, W., Hooker, C. A. (Eds.) *Beyond Ruling Reason: Non-formal Approaches to Rationality.*

Bickhard, M. H. (in preparation-a). From Epistemology to Rationality.

Bickhard, M. H. (in preparation-b). *The Whole Person: Toward a Naturalism of Persons*. Harvard.

Bickhard, M. H. (in preparation-c). Interaction and Representation.

Bickhard, M. H., Campbell, R. L. (1989). Interactivism and Genetic Epistemology. *Archives de Psychologie, 57*(221), 99-121.

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., Christopher, J. C. (1994). The Influence of Early Experience on Personality Development. *New Ideas in Psychology*, *12*(3), 229-252.

Bickhard, M. H., Richie, D. M. (1983). *On the Nature of Representation: A Case Study of James J. Gibson's Theory of Perception*. New York: Praeger.

Bickhard, M. H., Terveen, L. (1995). *Foundational Issues in Artificial Intelligence and Cognitive Science — Impasse and Solution*. Amsterdam: Elsevier Scientific.

Boolos, G. S., Jeffrey, R. C. (1989). *Computability and Logic. 3rd ed*. Cambridge.

Bourgeois, P. L., Rosenthal, S. B. (1983). *Thematic Studies in Phenomenology and Pragmatism*. Amsterdam: Grüner.

Brooks, R. A. (1991a). Intelligence without Representation. *Artificial Intelligence, 47*(1-3).

Brooks, R. A. (1991b). New Approaches to Robotics. *Science, 253*(5025), 1227-1232.

Campbell, D. T. (1974). Evolutionary Epistemology. In P. A. Schilpp (Ed.) *The Philosophy of Karl Popper*. LaSalle, IL: Open Court.

Campbell, D. T. (1990). Epistemological Roles for Selection Theory. In N. Rescher (Ed.) *Evolution, Cognition, and Realism*. Lanham, MD: University Press, pp. 1-19.

Campbell, R. L., Bickhard, M. H. (1986). *Knowing Levels and Developmental Stages.* Basel: Karger.

Campbell, R. L., Bickhard, M. H. (1987). A Deconstruction of Fodor's Anticonstructivism *Human Development, 30*(1), 48-59.

Campbell, R. L., Bickhard, M. H. (1992a). Clearing the Ground: Foundational Questions Once Again. *Journal of Pragmatics*, *17(5/6),* 557-602.

Campbell, R. L., Bickhard, M. H. (1992b). Types of Constraints on Development: An Interactivist Approach. *Developmental Review, 12*(3), 311-338.

Carlson, N. R. (1986). *Physiology of Behavior*. Boston: Allyn and Bacon.

Castañeda, H. (1989). *Thinking, Language, and Experience.* U. of Minnesota.

Chapman, D. and Agre, P. (1986). Abstract Reasoning as Emergent from Concrete Activity. In M. P. Georgeff, A. L. Lansky, (Eds.) *Reasoning about Actions and Plans, Proceedings of the 1986 Workshop*. Timberline, Oregon, 411-424.

Cherian, S., Troxell, W. O. (1995). Interactivism: A Functional Model of Representation for Behavior-Based Systems. In Morán, F., Moreno, A., Merelo, J. J., Chacón, P. *Advances in Artificial Life*: Proceedings of the Third European Conference on Artificial Life, Granada, Spain. (691-703). Berlin: Springer.

Cherian, S., Troxell, W. O. (in press). Intelligent behavior in machines emerging from a collection of interactive control structures. *Computational Intelligence*, 1994. Blackwell Publishers. Cambridge, Mass. and Oxford, UK.

Churchland, P. M. (1981). Eliminative Materialism and the Propositional Attitudes. J. *of Philosophy, 78*(2), 67-90.

Churchland, P. M. (1984). *Matter and Consciousness*. MIT.

Clancey, W. J. (1989). The Knowledge Level Reinterpreted: Modeling How Systems Interact. *Machine Learning* 4, 285-291.

Clancey, W. J. (1991). The Frame of Reference Problem in the Design of Intelligent Machines. In VanLehn, K., ed. *Architectures for Intelligence: The Twenty-Second Carnegie Symposium on Cognition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Clark, A. (1989). *Microcognition*. MIT.

Dennett, D. C. (1969). *Content and Consciousness*. London: Routledge & Kegan Paul.

Dennett, D. C. (1978). Intentional systems. In D. C. Dennett (Ed.) *Brainstorms*. Montgomery, Vt.: Bradford Books.

Dennett, D. C. (1987). *The Intentional Stance*. MIT.

Dennett, D. C. (1991). *Consciousness Explained*. Little, Brown.

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

Dretske, F. I. (1988). *Explaining Behavior*. Cambridge, MA: MIT.

Dreyfus, H. L. (1967). Why Computers Must Have Bodies in order to be Intelligent. *Review of Metaphysics, 21*, 13-32.

Dreyfus, H. L. (1982). Introduction. In H. L. Dreyfus (Ed.) *Husserl: Intentionality & Cognitive Science*. (1-27). MIT.

Dreyfus, H. L. (1991a). *Being-in-the-World*. MIT.

Dreyfus, H. L. (1991b). Heidegger's Hermeneutic Realism. In D. R. Hiley, J. F. Bohman, R. Shusterman (Eds.) *The Interpretive Turn.* (25-41) Cornell.

Dreyfus, H. L., Dreyfus, S. E. (1986). *Mind Over Machine*. New York: Free Press.

Dreyfus, H. L., Dreyfus, S. E. (1987). How to Stop Worrying about the Frame Problem even though its Computationally Insoluble. In Z. W. Pylyshyn (Ed.) *The Robot's Dilemma: The Frame Problem in Artificial Intelligence*. (95-111). Norwood, NJ: Ablex.

Eilenberg, S. (1974). *Automata, Languages, and Machines. Vol. A* New York: Academic.

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

Fodor, J. A. (1981). The present status of the innateness controversy. In J. Fodor, *RePresentations* . Cambridge: MIT Press (pp. 257-316).

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology.* Cambridge, Mass.: MIT.

Fodor, J. A. (1986). Why Paramecia don't have Mental Representations. In P. A. French, T. E. Uehling, H. K. Wettstein (Eds.) *Midwest Studies in Philosophy X: Studies in the Philosophy of Mind*. Minnesota, 3-23.

Fodor, J. A. (1987). *Psychosemantics*. Cambridge, MA: MIT Press.

Fodor, J. A. (1990). *A Theory of Content*. Cambridge, MA: MIT Press.

Ford, K. M., Hayes, P. J. (1991). *Reasoning Agents in a Dynamic World: The Frame Problem*. Greenwich, CT: JAI Press.

Gadamer, Hans-Georg (1975). *Truth and Method*. New York: Continuum.

Gadamer, Hans-Georg (1976). *Philosophical Hermeneutics*. Berkeley: University of California Press.

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*. Hillsdale, NJ: Erlbaum.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Ginzburg, A. (1968). *Algebraic theory of automata*. New York: Academic Press.

Glymour, C. (1988). Artificial Intelligence is Philosophy. In J. H. Fetzer (Ed.) *Aspects of Artificial Intelligence*. Kluwer Academic.

Guignon, C. B. (1983). *Heidegger and the Problem of Knowledge*. Indianapolis: Hackett.

Hahlweg, K., Hooker, C. A. (1989). *Issues in Evolutionary Epistemology*. SUNY.

Hanson, P. P. (1990). *Information, Language, and Cognition*. University of British Columbia Press.

Heidegger, M. (1962). *Being and Time*. New York: Harper & Row.

Hooker, C. A. (1995). *Reason, Regulation, and Realism: Towards a Regulatory Systems Theory of Reason and Evolutionary Epistemology*. SUNY.

Hooker, C. A. (in preparation). Toward a Naturalised Cognitive Science.

Hopcroft, J. E., Ullman, J. D. (1979). *Introduction to Automata Theory, Languages, and Computation*. Reading, MA: Addison-Wesley.

Houser, N., Kloesel, C. (1992). *The Essential Peirce. Vol. 1*. Indiana.

Kaplan, D. (1979a). On the logic of demonstratives. In P. French, T. Uehling, Jr., & H. Wettstein (Eds.), *Contemporary Perspectives in the Philosophy of Language*. Minneapolis: U. of Minnesota Press, 401-412.

Kaplan, D. (1979b). Dthat. In P. French, T. Uehling, Jr., & H. Wettstein (Eds.), *Contemporary Perspectives in the Philosophy of Language*. Minneapolis: U. of Minnesota Press, pp. 383-400.

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*. Oxford University Press, 481-563.

Kneale, W., & Kneale, M. (1986).*The development of logic.* Oxford: Clarendon.

Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). SOAR: An Architecture for General Intelligence. *Artificial Intelligence, 33*, 1-64.

Lewis, D. K. (1969). *Convention*. Cambridge, Mass.: Harvard.

Loewer, B., Rey, G. (1991). *Meaning in Mind: Fodor and his critics*. Oxford: Blackwell.

Loux, M. J. (1970). *Universals and Particulars*. Notre Dame.

Maes, P. (1990). *Designing Autonomous Agents*. MIT.

Malcolm, C. A., Smithers, T., Hallam, J. (1989). An Emerging Paradigm in Robot Architecture. In T Kanade, F.C.A. Groen, & L.O. Hertzberger (Eds.) *Proceedings of the Second Intelligent Autonomous Systems Conference*. (pp. 284-293). Amsterdam, 11--14 December 1989. Published by Stichting International Congress of Intelligent Autonomous Systems.

McCarthy, J., Hayes, P. (1969). Some Philosophical Problems from the Standpoint of Artificial Intelligence. In B. Meltzer, D. Michie (Eds.) *Machine Intelligence 4* (pp. 463-502). New York: American Elsevier.

McClelland, J. L., Rumelhart, D. E. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Vol. 2 Psychological and Biological Models*. Cambridge: MIT.

Millikan, R. (1984). *Language, Thought, and Other Biological Categories*. MIT.

Millikan, R. G. (1993). *White Queen Psychology and Other Essays for Alice*. MIT.

Moore, G. H. (1988). The emergence of first order logic. In W. Aspray & P. Kitcher (Eds.), *History and philosophy of modern mathematics* (pp. 95-135). Minneapolis: University of Minnesota Press.

Murphy, J. P. (1990). *Pragmatism*. Westview.

Newell, A. (1980a). Physical Symbol Systems. *Cognitive Science, 4*, 135-183.

Newell, A. (1980b). Reasoning, Problem Solving, and Decision Processes: The Problem Space as a Fundamental Category. In R. Nickerson (Ed.) *Attention and Performance VIII*. Hillsdale, NJ: Erlbaum.

Okrent, M. (1988). *Heidegger's Pragmatism*. Cornell.

Piaget, J. (1954). *The Construction of Reality in the Child*. New York: Basic.

Piaget, J., & Inhelder, B. (1973). *Memory and Intelligence*. New York: Basic.

Popper, K. (1965). *Conjectures and Refutations*. New York: Harper & Row.

Popper, K. (1972). *Objective Knowledge*. London: Oxford Press.

Pylyshyn, Z. (1987). *The Robot's Dilemma*. Ablex.

Radnitzky, G., Bartley, W. W. (1987). *Evolutionary Epistemology: Theory of Rationality, and the Sociology of Knowledge*. La Salle: Open Court.

Richard, M. (1983). Direct Reference and ascriptions of belief. *Journal of Philosophical Logic, 12*(4), 425-452.

Rosenthal, S. B. (1983). Meaning as Habit: Some Systematic Implications of Peirce's Pragmatism. In E. Freeman (Ed.) *The Relevance of Charles Peirce*. Monist, 312-327.

Rosenthal, S. B. (1986). *Speculative Pragmatism*. Open Court.

Rosenthal, S. B. (1987). Classical American Pragmatism: Key Themes and Phenomenological Dimensions. In R. S. Corrington, C. Hausman, T. M. Seebohm (Eds.) *Pragmatism Considers Phenomenology*. (37-57). Washington, D.C.: University Press.

Rosenthal, S. B. (1992). Pragmatism and the Reconstruction of Metaphysics: Toward a New Understanding of Foundations. In T. Rockmore, B. J. Singer *Antifoundationalism Old and New*. (165-188) Temple University Press.

Rumelhart, D. E. (1989). The Architecture of Mind: A Connectionist Approach. In M. I. Posner (Ed.) *Foundations of Cognitive Science*. MIT, 133-160.

Rumelhart, D. E., McClelland, J. L. (1986). *Parallel Distributed Processing. Vol. 1: Foundations*. MIT.

Santambrogio, M. (1990). Meinongian Theories of Generality. *Nous, XXIV*(5), 647-673.

Savage-Rumbaugh, E. S., Murphy, J., Sevcik, R. A., Brakke, K. E., Williams, S. L., Rumbaugh, D. M. (1993). Language Comprehension in Ape and Child. *Monographs of the Society for Research in Child Development, 58*(3-4), 1-221.

Schelling, T. C. (1963). *The strategy of conflict*. New York: Oxford University Press.

Shanon, B. (1991). Representations: Senses and Reasons. *Philosophical Psychology, 4*(3), 355-374.

Smith, B. C. (1987). *The Correspondence Continuum*. Stanford, CA: Center for the Study of Language and Information, CSLI-87-71.

Smithers, T. (1992). Taking Eliminative Materialism Seriously: A Methodology for Autonomous Systems Research. In Francisco J. Varela & Paul Bourgine (Eds.) *Toward a Practice of Autonomous Systems*. (pp. 31-40). MIT Press.

Thayer, H. S. (1973). *Meaning and Action*. Bobbs-Merrill.

Tomberlin, J. E. (1986). *Hector-Neri Castañeda*. Reidel.

Vuyk, R. (1981). *Piaget's Genetic Epistemology 1965-1980*. vol. II New York: Academic.

Waltz, D., Feldman, J. A. (1988). *Connectionist Models and Their Implications*. Norwood, NJ: Ablex.

Winograd, T., Flores, F. (1986). *Understanding Computers and Cognition*. Norwood, NJ: Ablex.