Chapter 2
Emergent Mental Phenomena

Mark H. Bickhard

2.1 Introduction

I argue that mental phenomena are emergent in biological phenomena, and, potentially, in artificial systems – though not with current technology. Developing this argument requires addressing several foundational issues. In particular, accounting for the normativity of mental phenomena requires a model of normative emergence, which, in turn, requires a model of genuine metaphysical emergence, which, so I argue, in turn requires an underlying process metaphysics. I will outline these framework preliminaries, in preparation for the discussion of emergent mental phenomena – with a focus on representing and consciousness – and some implications for the possibilities of artificial minds.

The emergence of normativity, given this background framework of process-emergence, occurs in certain kinds of thermodynamic systems. Within the general model of normative emergence, explicated in terms of normative *function*, I address, in turn, representational normativity, basic or primary consciousness, and reflective consciousness. Several conclusions follow from this overall model regarding the possibility of artificial mental systems, and for some other notions, such as that of 'uploading' of persons into computational frameworks.

I begin with an argument for process metaphysics – process metaphysics grounds the further levels of the model.

M. H. Bickhard (✉)
Department of Philosophy and Department of Psychology, Lehigh University,
Bethlehem, PA, USA
e-mail: mhb0@lehigh.edu

R. W. Clowes et al. (eds.), *The Mind-Technology Problem*, Studies in Brain and Mind 18, https://doi.org/10.1007/978-3-030-72644-7_2

2.2 Preliminaries

2.2.1 Metaphysics and Emergence

2.2.1.1 Why a Process Metaphysics?

There are several reasons for abandoning standard particle (or substance) metaphysics in favor of a process metaphysics.[1] These include conceptual problems, as well as problems with contemporary physics. For an example of a problematic conceptual problem, consider that a pure point particle model results in a world in which nothing ever happens – point particles have a zero probability of hitting each other. Furthermore, there is nothing to move such particles toward or away from each other, or for them to hold together if they were to ever be in other particles' vicinity. A particle model in which particle interactions occur via *fields* can partially resolve these issues,[2] but a field is a process – a field model is already a process model.

Furthermore, if we look to our best contemporary physics – quantum field theory – we find that there are no particles. Instead, there are various excitations of quantum fields (Fraser 2008; Halverson and Clifton 2002; Hobson 2013; Weinberg 1995).[3] Thus, we find fatal conceptual problems for a particle metaphysics, and no support within physics. Quantum field theory, on the other hand, is a clear process model.

2.2.1.2 Process Metaphysics and Emergence

Emergence is at best mysterious, if not impossible, within a substance or entity metaphysics. How can a new substance or entity 'emerge' from (some organization of) already existing substances or entities?[4] On the other hand, it is fairly easy to see how a new property could be instantiated in a new organization: an organization *is* a property of what is so organized. But, even if such a "new" property were dynamic (not just static), it is difficult to see how it could be a causally efficacious constituent of the world. One of Kim's arguments against emergence makes this clear.

[1] Bickhard 2009.

[2] Fields can introduce attractive and repulsive forces.

[3] These excitations and their interactions are quantized in the sense of involving whole integer or half integer amounts. This quantization is the last remnant of 'particle' intuitions, but it is akin to the whole or half integer wavelengths in the vibrations of a guitar string (or, in the case of half integer, perhaps a rope that is free at one end) – and there are no guitar sound (or rope) particles either.

[4] Parmenides argued against change, including emergence, and Democritean atoms and Empedoclean substances were proposed as satisfying this prohibition of fundamental metaphysical change – they can reconfigure and remix, but they do not themselves change. Aristotle accepted this prohibition of fundamental change, and the presupposition has dominated Western thought since then (Gill 1989; Campbell 2015; Bickhard 2009).

In the 'causal regularity' argument, Kim points out that a new configuration of particles might reliably yield a regularity of consequences, including causal consequences, but any such causal consequences would be due solely to the interactions among the particles that instantiated the configuration in the first place (Kim 1991). Configurations can yield causal regularities, but not causal powers.

This argument reveals rather clearly that the underlying assumption is that only particles can bear causal power. Configuration is not a particle, or an entity, or a substance, so it is not even a candidate for having causal power.[5] Configuration is just initial conditions, or boundary conditions, for the real causal interactions among the particles – configuration is just stage setting for the real causal dance. But emergence is supposed to be emergence from or within configuration (organization), so, in this view, it is ruled out *by assumption* that there could possibly be any new causal power emergent in organization. Note that this formally begs the question concerning emergent causal power: the very possibility is assumed to not exist.

In a process metaphysics, however, this default assumption about the locus of causality is flipped: processes are intrinsically organized, and processes influence other processes in crucial part due to their organization. Organization *cannot be delegitimated* as a potential locus of 'causal' influence without eliminating 'causality' entirely – without eliminating 'causality' from the world.[6] But, if organization can manifest 'causal' influence ('power'), then 'new' organization can manifest 'new' – emergent – efficacious influence in the world.

2.2.2 Normative Emergence, Function, Representation

2.2.2.1 Normative Emergence

If emergence per se is a metaphysical possibility, then perhaps the emergence of *normativity* is possible. There are long standing reasons why this should *not* be possible, the first of which is 'simply' that emergence itself is rather difficult to make any sense of within a classic particle or substance model. In addition to that problem, there is also the point that the substance/particle metaphysical world simply has no place, emergent or not, for normativity: it is a metaphysics of fact and cause. This split is enshrined in contemporary thought via (among other bases) Hume's 'argument' against being able to derive norms from facts.

But Hume's 'argument' is (arguably) unsound, and a shift to a process metaphysics not only makes emergence more generally possible, but also opens the possibility of accounting for normative emergence.

[5] For analyses of Kim's more well known argument – the pre-emption argument – showing that it too depends on the same underlying particle assumptions, see (Bickhard 2009, 2015).

[6] The scare quotes are because this kind of ongoing (coupling constant) influence among quantum fields (for example) does not fit well with standard causal chain models of causality.

2.2.2.2 Hume's Argument

Hume did not detail an argument for his "no ought from is" maxim, but claimed that it "seems altogether inconceivable, how this new [normative] relation can be a deduction from others, which are entirely different from it" (see Hume 1978, Book III. Part I. Section I. 469–470). Hume's unstated 'argument' has generated considerable work attempting to explicate and formalize it (e.g., Schurz 1997); I offer an explication of Hume's point that shows it to be valid but unsound.[7] In being unsound, Hume's maxim ceases to be a barrier to the possibility of a model of normative emergence.

A central aspect of deduction is that of definition; e.g., in deducing theorems about triangles from Euclid's axioms, a definition of 'triangle' based on the terms in the axioms (e.g., point and line) is required. Hume's maxim is readily derived if we consider how any new (e.g., normative) terms could be validly introduced in deductions from strictly factual premises. Any new terms must be defined making use of terms already available. These may include those introduced by prior definitions, and those may make use of still prior definitions, but all such definitions must ultimately be in terms that are originally available in the premises. But, given such a hierarchy of definitions, each 'new' term can be back-translated via its definition into prior terms,[8] again through the layers of the hierarchy, till all terms in the conclusion are unpacked into terms in the premises. But, by assumption, those premise terms are all factual, so any valid conclusions must also be strictly factual – not normative. And we have Hume's 'inconceivability' of deduction.

This argument is valid, but it is based on the unstated premise that all definitions are 'abbreviatory' – that all definitions can be back-translated through. But that premise is false, and, if so, the argument is unsound. There was no alternative to abbreviatory, back-translation, definition in Hume's time, but in the nineteenth century *implicit definition* was introduced, with Hilbert being one of its major proponents.[9]

The intuition of implicit definition is that a system of relations – an axiom system for geometry for example – implicitly defines the class of all of its models. It implicitly defines the class of all of the ways in which it can be interpreted that honor all of the relations. Two Xs determine a Y, for example, can be interpreted as "two points determine a line", but it also turns out that it can be interpreted as "two lines determine a point" (the intersection of the lines, so long as intersections at infinity are considered). Implicit definition can also be understood non-formally (Hale and Wright 2000) and also dynamically.[10]

---

[7] As well as that it is related to more general anti-emergentist arguments. See Bickhard (2009, 2015).

[8] I.e., substitute the defining term, phrase, or clause for the defined term. The defined terms are abbreviations of the definiens, so 'fill out' all of the abbreviations.

[9] E.g., Hilbert developed an implicit definitional approach to geometry (Hilbert 1971). See Chang and Keisler (1990) for a modern formal development of implicit definition.

[10] The model of functional presupposition, developed later in the text, is an example of dynamic implicit definition.

The key point for current purposes is that implicit definition is a powerful form of definition (e.g., model theory is based on it) and it is not abbreviatory – it does not permit back-translation. The (re-constructed) Hume argument, therefore, cannot go through, and the "no ought from is" maxim does not necessarily hold. The overall argument is unsound: it involves a false premise concerning definition.

The *barrier* of the Humean maxim is, thus, cleared. But that does not constitute a *model* of normative emergence. I turn to that now.

2.2.2.3 Normative Function

To this point, the discussion has been 'brush clearing' – clearing apparent barriers to emergence and to normative emergence. I now turn to a model of the emergence of normativity in the form of normative function – the sense of function in which it makes sense to distinguish function from dysfunction, as in "This kidney is dysfunctional."

The model of function is grounded on a crucial asymmetry between thermodynamically differing kinds of processes – in particular, between process organizations that are stable in energy wells and those that are (relatively) stable in far from thermodynamic equilibrium conditions. Processes are always ongoingly changing, but some *organizations* of process can remain stable for some time as organizations, and that stability can be of (at least) two differing kinds.

Thus, there are two kinds of such stability that will be of concern here: The first kind are process organizations that remain stable because they are in an 'energy well'. Such organizations will remain stable unless and until some above threshold amount of energy impinges on them that is sufficient to disorganize them – to knock them out of the energy well.

An atom would be a canonical example. It is a furious process of quantum fields that can remain stable for cosmological time periods, if not disrupted. One crucial feature of energy well stabilities is that, if they are isolated, they go to thermodynamic equilibrium and remain in their organization indefinitely.

In contrast are far from equilibrium stabilities of process. Like energy well stabilities, far from equilibrium organizational stabilities can persist for some time. Unlike energy well stabilities, however, they cannot be isolated: being far from equilibrium is a *relational* condition that must be *maintained*. If isolated, they go to equilibrium and the process organization ceases to exist.

A canonical example of this would be a candle flame. If isolated, the flame ceases: if isolated, the process goes to equilibrium and is, thus, no longer far from equilibrium. Far from equilibrium conditions must be maintained.

A candle flame also illustrates a further property: it makes contributions to its own stability. The flame maintains above combustion threshold temperature, vaporizes wax in the wick, melts wax in the candle, and induces convection, which brings in oxygen and removes waste products. It contributes to its own stability in several ways; it is *self-maintaining*.

There is an additional property that a candle flame does not have, but a bacterium does. If a candle flame is running out of candle, it cannot change the process in any way to adapt. A bacterium, in contrast, will tend to swim upward in a sugar gradient, which is a contribution to its stability, but, if it is going toward lower sugar concentrations in the gradient, it will tend to tumble, and then resume swimming.[11] Swimming contributes to stability in some conditions (up a gradient) and not in others, and the bacterium can adjust what it is doing in order to maintain the property of contributing to its own persistence. It self maintains its condition of being self maintenant: it is *recursively self maintaining*.

The crucial point here is that contributions to the maintenance of a far from equilibrium process are *contributory* – they are useful, *functional*, *for* and *relative to* the persistence of that organization of process. This is a normative relationship: it can hold or not hold, and it makes a difference to the system whether it holds or not.[12]

The structure of the model of emergent function, thus, is that:

(1) The asymmetry between energy well and far from equilibrium processes yields
(2) An asymmetry between contributions to the thermodynamic maintenance of far from equilibrium processes and impairments to that maintenance, which, in turn, grounds
(3) The emergent asymmetry between functional and dysfunctional.

The further properties of being self-maintenant and recursively self-maintenant, in turn, begin a hierarchy[13] of more complex forms of autonomy of far from equilibrium systems. Note that this sense of autonomy focuses on the interdependence between a system and its environment – the ability of the system to adjust itself *and* its environment toward functionality for the system – rather than autonomy in the sense of independence or freedom *from* the environment.

2.2.2.4 Representational Truth Value

The normativity of function grounds a further emergent normativity – the normativity of representational truth value. This emergence occurs with respect to a particular function that is necessary for any agent interacting with its world: the function of being able to select what to do next, or to guide ongoing interaction, on the basis (among other things) of what the possible interactions might be in the current situation. That is, there must be some (functional) indications of, anticipations of, further courses of possible interaction in the current situation among which the system (organism) can select.

---

[11] See Campbell (1974). Bacteria can be more complex than this, but the simple example illustrates the point.

[12] The sense in which this is a functionality relative to a system can be illustrated with the case of the beating heart of a parasite, which is functional for the parasite but dysfunctional for the host.

[13] Or, more complexly, a lattice or weave.

A frog, for example, might have indications that it could flick its tongue in any of several directions and eat. An external observer might see a couple of flies and a worm in those directions. Such indications are anticipatory that, if selected, the interaction would proceed as indicated. But such anticipations can be in error: they can be false. Representational truth value is emergent in such anticipations. Truth value bearing anticipatory indication, in turn, is the basis for representing in general, including more complex representing.[14]

2.2.2.5 Content

Furthermore, an indication of the potentiality of an interaction *presupposes* that sufficient supporting conditions for that interaction to succeed hold in the environment – such as, perhaps, a fly or worm. If the anticipatory indication fails, then those supporting conditions did not hold. This implicit presupposition of supporting conditions in the environment, thus, is *about* that environment. Presupposed supporting conditions constitute a model of *content* – the supporting conditions are what are *supposed* to exist in order for the anticipation to hold. Content in this sense, however, is implicit, not explicit (Bickhard 2009).

2.2.2.6 Complex Representing

The model of representing in terms of anticipations of potential interacting has two important resources for modeling more complex representing. The first has already been indicated: the frog has *branching* indications, hopefully triggered by, for example, flies or worms.[15] The second resource is that such indications can iterate. Again, the frog can provide an example: perhaps among the frog's indications are that, if it were to move a little to the left, another pair of worms would come into range. So, indications may be that, some interaction is possible, and, if it were engaged in, it would yield the conditions for further interactions.

Branching and iterating indications of interactive potentialities can link to create potentially complex webs of interactive anticipation. In humans, these webs are vast. I have dubbed such webs as *situation knowledge* – interactive knowledge of what the organism could do in a broad (branched and iterated) sense.

[14] This model involves a shift in what is taken to be most centrally criterial for representing. Standard models assume that the crucial property of representing is that of having some sort of denotational or referential relationship with the environment – some sort of critical contact with or correspondence with the environment. In the model outlined above, the criterial property for representing is that of bearing (potential) truth value. Contact with the environment is also centrally important, but, contrary to standard assumptions, contact per se does not yield truth value (Bickhard 2009; Oguz and Bickhard 2018).

[15] "Hopefully" because, for example, they might also be triggered by (contact with) a tossed pebble, in which case they would be false.

Situation knowledge is not constant. It is ongoingly changing on the basis of processes occurring in the environment and on the basis of activities of the organism. Situation knowledge is undergoing constant maintenance and updating. Such processes of keeping situation knowledge up to date is *apperception*.

The situation knowledge web, in turn, is a realm in which more canonical forms of representing can be modeled. For example, a child's toy block will support a (sub-)web of possible interactions that is *closed* in the sense that any manipulatory or perceptual interaction with the block can connect with, lead to, any other such interaction with the block via some intermediary interactions, such as rotations of the block. Further, this internally reachable subweb is itself invariant under many other kinds of processes, such as throwing the block, leaving it on the floor, putting it away in the toy box, and so on. It is *not* invariant under all interactions, however: burning the block, for example, eliminates the support for that situation knowledge subweb. This model of representing a small manipulable object is basically Piaget's model of representing a small object stated within the terms of this model (Piaget 1954). The Piaget model can be borrowed from in this manner because both are 'pragmatic' models, based on action and interaction rather than on correspondence.[16]

2.3 Consciousness: Primary and Reflective

The model of representing supports a model of consciousness with two aspects, primary and reflective. These aspects do not have to occur together, though reflective cannot occur without primary – primary consciousness is what reflective consciousness can reflect upon. But primary forms of consciousness can and do occur without reflective, e.g., in some species and in neonate humans. If this model is correct, consciousness is not a unitary kind of phenomena.

2.3.1 Primary Consciousness

In particular, the model of interaction, situation knowledge, and apperception already outlined provides an account of a process flow that is intrinsically contentful – the contents of situation knowledge, the apperceptive processes that maintain it, and the anticipatory processes that make use of it. This is a flow of content in the sense of anticipations of possibility and relationships among them, not in terms of encoding correspondences (or denotations). It is an intentional flow in terms of the differentiations of the world induced by interactions with that world, and the

[16] This model borrows from Piaget for other phenomena as well, though usually with modifications: Piaget evidences what I take to be errors in some aspects and parts of his model (Bickhard 1988; Bickhard and Campbell 1989).

anticipatory indicative relations among them. It is a partition epistemology (Levine 2009) – partitions induced by differentiations – rather than a correspondence epistemology.[17]

Furthermore, this flow is intrinsically embodied – a body is necessary for *inter*-action. Consequently, it is also situated and from a point of view. The model, thus, accounts for multiple properties of consciousness, properties that arguably exist in simple organisms as well as human beings.[18]

But this does not model all properties of human consciousness (and perhaps some other higher primates). In particular, it does not model conscious reflection, or *reflective consciousness*.

2.3.2 Reflective Consciousness

Primary consciousness involves a taking into consideration the agent's relationships with the world and with the potentialities of that world. It is a kind of *awareness* of the world. But primary consciousness does not offer an account of awareness of awareness, of *reflective consciousness*. Interacting is asymmetric; it is normatively 'about'; it involves a normative agent and a world. In particular, interacting, thus awareness, is not in itself reflective.

But reflection can clearly occur: any reflection on the issue is an instance of the phenomenon. The interactivist model argues for a level of *awareness of* processes of *awareness* (interaction, situation knowledge, apperception) that has evolved (in various forms and degrees) in the CNS of some species (Bickhard 2015a, b).

Reflective consciousness, thus, is reflective awareness of processes of awareness. It involves differentiated aspects of the CNS; it requires primary consciousness to reflect upon; but primary consciousness can exist in various species and individuals without the possibility of reflection: consciousness is not a unitary kind of phenomena (Bickhard 2005). Reflection is constituted as a second level interacting with the first level, which, in turn, interacts with the environment.[19]

---

[17] In general, cognitively simpler organisms will involve less complex situation knowledge webs and more general differentiations – e.g., the differentiation of a "keep swimming" condition.

[18] For consideration of phenomena such as emotion, motivation, and other psychological phenomena, see (Bickhard 2000, 2003).

[19] Note the partial convergence with HOT theories of consciousness (Lau and Rosenthal 2011; Rosenthal 2010).

2.3.3 Experiencing

Primary conscious flow is a strong candidate for modeling *experiencing*. It involves experiencing qualities or properties of interacting with the world. Reflective consciousness, in turn, involves experiencing the properties or qualities of primary experiencing – experiencing experiencing.

These properties of the experiencing of experiencing are commonly reified into discrete elements called qualia. This is incorrect: experiencing is a flow. But, worse, qualia are often taken to not only be (discretized) experiencing of experiencing, but also to be *constitutive* of *basic* experiencing – as in sense data models. This would make qualia the experiencing of qualia. This is a tight metaphysical circle that makes understanding experiencing ultimately impossible (Bickhard 2005).

Qualia (overlooking the assumed discreteness) are *results of reflection*, not constituents of what is reflected on.[20] Reflective experiencing is much easier to understand if this is taken into account. Furthermore, what are at times taken to be the "easy" problems – e.g., normative representing – are much harder to understand than often considered – they involve, for example, the emergence of normativity. The realm of consciousness and experience looks significantly different when viewed in these interactive terms.

2.3.4 An Argument Against the Possibility of the Emergence of Conscious Experience

There are strong arguments in the literature against any such possibility of the emergence of consciousness and experiencing. As with the problems induced by particle metaphysics and by Hume's "argument", I will address one class of these arguments with an intent of brush clearing (again) – showing that the arguments are unsound, in that they make an underlying false assumption.

The argument that I wish to address has the following general form: We can model numerous "easy" problems, such as representation, in terms of causal functionalism – in terms of, for example, symbolic or connectionist encodings. But causal function is 'just' a standard causal relation that is picked out of a realm of causal consequences as being relevant to some consideration, such as constituting part of what makes a computer. A transistor, for example, has multiple causal consequences, such as creating heat, but the only one that counts as causally functional is the (perhaps) switching function that it introduces in a circuit.

But causality is indifferent to consciousness and experience. A causally identical system or organism could have wildly different experiences, or none at all (a zombie). Possibilities of inverted qualia, dancing qualia, and so on are perfectly consistent with whatever causal functional processes make up a person, so that experiential

[20] This is basically Dewey's criticism of Russell's sense data model (Dewey 1960; Tiles 1990).

realm is independent of those causal functional processes: the experiencing could be wildly different, or not exist at all.

Causal functionality seems adequate to many ‘mental’ phenomena – those are ‘easy’ problems – but cannot be adequate to the qualia of experience – that is the ‘hard’ problem (e.g., Chalmers 1996).

One crucial assumption in this argument is manifest: every functional relationship is ‘just’ a selected causal relationship. This ignores, for example, the possibility of normative function. It might be argued, however, that even normative function is still ‘just’ cause, just cause selected by some sort of *intrinsic* functionality, so it is still intrinsically ‘just’ cause and therefore cannot account for the qualia of experience.

That line of dispute can be continued further, but there is a deeper problem that I wish to point out that undercuts that framework of issues: there is an assumption in this argument, including in its reliance on ‘cause’, that all (crucial) relationships are *external*.

The distinction between metaphysical internal and external relationships is mostly lost in contemporary philosophy. It was important at the turn into the twentieth century, and for decades after, but Russell tried to reject internal relations and Quine pretty successfully did so.

The distinction is between relations that are in some basic sense essential to something’s being what it is (internal) and some relations that are irrelevant to what something is. An external relationship might be between an effect and its cause: it would be that effect even if from a quite different cause. An example of an internal relationship might be that between an arc of a circle and the point that is the center of that circle: it could not be an arc of that circle if it did not have that relation to that point.

Note that a background assumption of a particle metaphysics in which particles are independent particulars is a framework in which the basic level of existence is composed of entities that cannot have internal relations (if they did have internal relations, they would not be independent particulars, thus not particles; Seibt 1996, 2009, 2010, 2012; Campbell 2015).

Similarly, a mechanistic, causal functional framework, is one in which all relations are external (cause is a classical example of an external relation) *by assumption*. The case of the presuppositions of normatively functional anticipations, however, is a kind of *internal* relation: in no universe with thermodynamics like this one could there exist normative anticipations of possible interactions that did not have (was not related to) presuppositions of the possible supports for those anticipations.

Even if this model of presupposition is not accepted, the background assumption in the “hard” problem arguments of external relations ignores the very possibility of internal relations, and, thus, the arguments are unsound: if the relations between system processes and conscious experience are internal, then zombies, inverted and dancing qualia, and so on, are impossible. So the assumption that the relations are external has to be established in order for these arguments to go through, but that assumption is not even addressed. That suffices for the 'brush clearing' of these arguments: they are not a barrier because they are unsound.

2.4 Artificial?

The model outlined is *not* a computational model.[21] As such, it significantly alters the questions concerning the possibility of artificial intelligence, agency, cognition, and consciousness: If this model, or anything like it, is correct, then none of these phenomena can be constituted as strictly computational systems.

The model, however, does not preclude the possibility of artificial systems that have emergent mental properties. But this cannot be done with contemporary (mechanistic/computational) technology. Instead, the emergence of normativity, including in its myriad particular instances, requires particular kinds of far from thermodynamic equilibrium systems, not (just) mechanistic systems (Bickhard 2007).

And, though not impossible in principle, creating artificial complex recursively self-maintenant systems requires creating something that constitutes life (at least metabolically). A major challenge.

2.4.1 Uploading?

One consequence of the shift in this model to a continuous dynamic framework is the impossibility of what has been called "uploading" – the uploading of a person into a computational condition.

Computation requires special conditions to be imposed and maintained on underlying continuous processes. In particular, the physical differences that underlie what are called ones and zeros must be maintained as detectably distinct and must be converted consistently across all physical substrates involved in the computational processes: that is, magnetization up and down must be kept distinct, and must be converted consistently across categories of high voltage/low voltage, high electromagnetic pulse/low pulse, high charge/low charge, spin up/spin down, light pulse/no light pulse, and so on. Maintaining such differences and the consistencies of conversions among them is required in order for computation on such categories (e.g., ones and zeros) to be possible, but such maintenance is in the face of physical dynamic tendencies toward smearing and erasure of the distinctions. Computers are good at this; otherwise they wouldn't be consistent, or would fail altogether.

[21] It is, in fact, *anti*-computational. Furthermore, for similar reasons of fundamental incoherence in models of representing (as well as other phenomena), it is also anti-connectionist – not with respect to the technologies per se, but with respect to claims that they might capture the nature of representing. For discussions of this and some other models and positions in the literature, see (Bickhard 2009, 2014, 2016a, b; Bickhard and Terveen 1995).

The very notion of uploading, however, depends on a background assumption of such discrete causal functionalism. If the model outlined above is correct, then there are two fatal problems with this assumption: (1) *causal* functionalism does not suffice: causal functionalism is a mechanistic model and cannot realize the necessary normativities, and (2) real dynamic phenomena cannot be captured by discrete approximations. For example, chaotic processes can be useful to an organism for the sake of unpredictability-in-principle in situations of competition or predator/prey interactions – it can be unadaptive to be predictable, and chaos makes predictability impossible. This is because chaotic processes are sensitive to close to infinitesimal differences in (initial and/or boundary) state and no process can measure such fine differences, so no reliable prediction can be made. So, insofar as anything like chaos is important in the realization of mental processes, no possible measurement is capable of capturing the crucial differences in state. So, no uploading can occur.

Furthermore, emergence is important, according to this model, not only in the broad sense of normative, functional, representing, and other kinds of emergence, but also in ongoing processes, such as emergent variations in mental variation and selection problem solving processes (Bickhard 2002), and, again, this cannot be captured in discrete measurements or discrete systems: (1) it is likely to itself be chaotic, and (2) the self-organizational properties of irreversible processes is involved (Bickhard, 2002), and this too involves fine sensitivity to initial and boundary conditions.

The very notion of uploading, thus, depends on the background assumption of the adequacy of computational models to mental phenomena. The interactivist model precludes that adequacy, and thus precludes the possibility of uploading.

2.5 Conclusion

Artificial systems with emergent mental properties are possible in principle, but not with current technology. Appearances that current information processing – computational, connectionist, and so on – technology might be adequate are false. What is required are particular kinds of far from thermodynamic equilibrium processes that go beyond 'mechanistic' causal functional models and technologies.

I have outlined a model of the emergence of normative phenomena, particularly normative function, representing, and experiencing in this chapter, and investigated some of the consequences for the possibility of artificial experiencers. The very possibility of such emergences, in turn, depends on a process metaphysics, which enables a model of metaphysical emergence, which enables a model of normative emergence.

Thus, the framework of process, emergence, normative emergence, and the specific models of various kinds of normative emergence form an integrated conceptual whole. No one part of it stands alone. Within this framework, the creation of artificial experiencing systems is possible, but only with a technology that can design and create complex recursively self-maintaining far from thermodynamic equilibrium systems.

References

Bickhard, M. H. (1988). Piaget on variation and selection models: Structuralism, logical necessity, and interactivism. *Human Development, 31*, 274–312.

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

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

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

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

Bickhard, M. H. (2007). Mechanism is not enough. In Q. Gonzalez, M. Eunice, W. F. G. Haselager, I. E. Dror (Eds.) Mechanicism and autonomy: What can robotics teach us about human cognition and action? (Special issue of pragmatics and cognition, Vol. 15, issue 3, pp. 573–585). Amsterdam: Benjamins

Bickhard, M. H. (2009). The interactivist model. *Synthese, 166*(3), 547–591.

Bickhard, M. H. (2014). What could cognition be, if not computation … or connectionism, or dynamic systems? *Journal of Theoretical and Philosophical Psychology, 35*(1), 53–66.

Bickhard, M. H. (2015). The metaphysics of emergence. *Kairos, 12*, 7–25.

Bickhard, M. H. (2015a). Toward a model of functional brain processes I: Central nervous system functional micro-architecture. *Axiomathes, 25*(3), 217–238.

Bickhard, M. H. (2015b). Toward a model of functional brain processes II: Central nervous system functional macro-architecture. *Axiomathes, 25*(4), 377–407.

Bickhard, M. H. (2016a). The anticipatory brain: Two approaches. In V. C. Müller (Ed.), *Fundamental issues of artificial intelligence* (pp. 259–281). Cham: Springer.

Bickhard, M. H. (2016b). Inter- and En- activism: Some thoughts and comparisons. *New Ideas in Psychology, 41*, 23–32.

Bickhard, M. H., & Campbell, R. L. (1989). Interactivism and genetic epistemology. *Archives de Psychologie, 57*, 99–121.

Bickhard, M. H., & Terveen, L. (1995). *Foundational issues in artificial intelligence and cognitive science: Impasse and solution*. Amsterdam: Elsevier Scientific.

Campbell, D. T. (1974). Evolutionary epistemology. In P. A. Schilpp (Ed.), *The philosophy of Karl Popper* (pp. 413–463). LaSalle: Open Court.

Campbell, R. (2015). *The metaphysics of emergence*. New York: Palgrave Macmillan.

Chalmers, D. J. (1996). *The conscious mind*. Oxford: Oxford University Press.

Chang, C. C., & Keisler, H. J. (1990). *Model theory*. North Holland.

Dewey, J. (1960/1929). *The quest for certainty*. New York: Capricorn Books.

Fraser, D. (2008). The fate of "particles" in quantum field theories with interactions. *Studies in History and Philosophy of Modern Physics, 39*, 841–859.

Gill, M.-L. (1989). *Aristotle on substance*. Princeton: Princeton University Press.

Hale, B., & Wright, C. (2000). Implicit definition and the a priori. In P. Boghossian & C. Peacocke (Eds.), *New essays on the a priori* (pp. 286–319). Oxford: Oxford University Press.

Halvorson, H., & Clifton, R. (2002). No place for particles in relativistic quantum theories? *Philosophy of Science, 69*(1), 1–28.

Hilbert, D. (1971). *The foundations of geometry*. La Salle: Open Court.

Hobson, A. (2013). There are no particles, there are only fields. *American Journal of Physics, 81*, 211. https://doi.org/10.1119/1.4789885.
Hume, D. (1978). *A treatise of human nature*. Index by L. A. Selby-Bigge; notes by P. H. Nidditch. Oxford: Oxford University Press.
Kim, J. (1991). Epiphenomenal and supervenient causation. In D. M. Rosenthal (Ed.), *The nature of mind* (pp. 257–265). Oxford: Oxford University press.
Lau, H., & Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends in Cognitive Sciences, 15*(8), 365–373.
Levine, A. (2009). Partition epistemology and arguments from analogy. *Synthese, 166*(3), 593–600.
Oguz, E., & Bickhard, M. H. (2018). Representing is something that we do, not a structure that we "use": Reply to Gładziejewski. *New Ideas in Psychology, 49*, 27–37.
Piaget, J. (1954). *The construction of reality in the child*. New York: Basic.
Rosenthal, D. (2010). How to think about mental qualities. *Philosophical Issues, 20*, 368–393.
Schurz, G. (1997). *The is-ought problem: An investigation in philosophical logic* (Trends in logic, Vol. 1). Dordrecht: Kluwer Academic.
Seibt, J. (1996). The myth of substance and the fallacy of misplaced concreteness. *Acta Analytica, 15*, 119–139.
Seibt, J. (2009). Forms of emergent interaction in general process theory. *Synthese, 166*(3), 479–512.
Seibt, J. (2010). Particulars. In R. Poli & J. Seibt (Eds.), *Theories and applications of ontology: Philosophical perspectives* (Vol. 1, pp. 23–57). New York: Springer.
Seibt, J. (2012). Process philosophy. In *Stanford encyclopedia of philosophy.* http://plato.stanford.edu/entries/process-philosophy/
Tiles, J. E. (1990). *Dewey*. London: Routledge.
Weinberg, S. (1995). *The quantum theory of fields. Vol. 1. Foundations*. Cambridge: Cambridge University Press.

Mark H. Bickhard is the Henry R. Luce Professor in Cognitive Robotics and the Philosophy of Knowledge at Lehigh University. He is affiliated with the Departments of Philosophy and Psychology, and is director of the Institute for Interactivist Studies. His work ranges from process metaphysics and emergence to consciousness, cognition, language, and functional models of brain processes to persons and social ontologies.