Abstract:
This paper assesses the main argument of Thomas Nagel’s recent book, *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature Is Almost Certainly False*. The paper agrees with Nagel that, as an approach to the relation between mind and matter and the mystery of subjective experience, neutral monism is more likely to be true than either materialism or idealism. It disagrees with Nagel by favoring a version of neutral monism based on emergence rather than on a reductive pan-psychism. However, the paper invokes a reductive view when applied to information (as opposed to psyche), and posits a hierarchy of types of information that span the domains of matter, life, and mind. Subjective experience is emergent, but also continuous with informational phenomena at lower levels.

Key words: reductionism, mind-body problem, subjective experience, emergence, autopoiesis, information, pan-psychism, neutral monism, materialism
Nagel’s argument

Thomas Nagel in *Mind & Cosmos* (2002) asserts that psychophysical reduction – explaining mind in terms of matter – has been and will forever be a failure, and that therefore “the materialist theory of evolution is almost certainly wrong.” “Mind” here refers to consciousness; more precisely, to subjective experience. Nagel briefly reviews various approaches that have been taken towards the relation of mind to matter (Table 1). He rejects the currently dominant approach of materialism and favors the alternative of neutral monism; specifically, its reductive – as opposed to emergent – version. Nagel thinks that materialist neo-Darwinism is also unable to explain rationality or justify value realism, but these extensions of his argument will not be considered here.

Table 1. Points of view about mind and matter
The dominant view is shown in italics; Nagel’s preferred view is shown in bold.

<table>
<thead>
<tr>
<th>Dualism</th>
<th>substance or attribute dualism</th>
<th>Monism</th>
<th>reductionism (eliminative or identity-theory)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>idealism</td>
<td>theism, absolute idealism, subjective idealism</td>
</tr>
<tr>
<td>neutral monism</td>
<td></td>
<td>reductive</td>
<td>emergent</td>
</tr>
</tbody>
</table>

While opposing materialism, Nagel also rejects idealism, its traditional alternative, which either invokes God as an ultimate mental reality (theism), takes ideas as more real than matter (absolute idealism), or grounds reality in human experience (logical positivism, phenomenology, constructivism). Idealism is rejected out of hand; Nagel concerns himself only with arguing for the inadequacy of materialism, specifically materialist neo-Darwinism (MND), for explaining consciousness. Nagel insists that this inadequacy is not a ‘local’ failure. Human (and animal) subjective experience is a salient manifestation of life, and failure to explain such experience is not a minor flaw. Nagel argues that since
extension of scientific theory to encompass consciousness will require radical change, current theory must be “false.” His objections to MND go even further. Aside from the problem of subjective experience, he doubts that evolutionary theory adequately explains the history – especially the origins – of life even in its simplest forms, where subjectivity is usually assumed to be absent. While this objection to evolutionary theory is a supplementary argument not dependent on Nagel’s primary thesis, it is clear that Nagel’s doubts about the adequacy of MND to account for the origin of life reinforces his doubts about its adequacy to encompass mind.

Rejecting both extremes of materialism and idealism, Nagel advocates neutral monism, and, between its reductive and emergent varieties, prefers the former. In reductive monism – different from materialist reductionism – entities have both material and mental properties. Mental properties are only discernable in certain complex entities, but along with physical properties, they go ‘all the way down.’ Reductive monism implies pan-psychism. In emergent neutral monism, by contrast, psychic properties only emerge at a certain level, being altogether absent at lower levels.

For Nagel, reductive and emergent monism each has its appeal, but each also has weaknesses. The weaknesses of reductive monism are that it attributes to simple entities psychic properties that so far are unobservable, that it is hard to conceptualize whole-part relationships involving mental properties, and that it seems to require not only mind but also life to go all the way down, a violation of Okham’s principle of parsimony. The weakness of emergent neutral monism is that the sudden appearance of psyche without precursors appears magical.

Nagel’s analyses include a synchronic “constitutive” mode of explanation and a diachronic “historical” mode, the completeness of explanation of any phenomenon requiring both modes. This explanatory dyad can be expanded into a structure-function-history triad (Gerard 1958), where structure refers to the internal order of a system, function to its participation in an external order, and history to qualitative change in both internal and external orders. In terms of this triad, constitutive analysis of mind includes its relation to brain (structure) and to behavior, language, and social communication (function), while historical analysis takes up the origin of mind and its evolutionary development. Note that ‘function’ in this triad is different than ‘function’ in functionalist theories of mind. In the structure-function-history triad, ‘function’ means external as opposed to internal; in functionalist explanations of mind, ‘function’ refers to form as opposed to substance.

Response to Nagel

**Emergent novelty need not be irreducible**

What follows is a brief for subjective experience as an emergent phenomenon, but one resembling other informational phenomena that occur in living systems. In Nagel’s terminology, this posits an emergent neutral monism, augmented by a reductive view. In this view, mind is both different from and similar to what evolutionarily preceded it.
Difference, i.e., newness, manifests emergence, while similarity, i.e., continuity with antecedents, supports the reductive view.

Nagel favors the reductive view because he thinks it unlikely that a completely new property can emerge without being present at lower levels. This assumption, that the properties of a whole must also be properties of its parts, is a “fallacy of decomposition.” A hydrogen molecule has vibrational modes of oscillation, but its constituent atoms do not. Water has fluidity, but individual water molecules, and their hydrogen and oxygen atoms, and their protons, neutrons, and electrons, and their quarks and leptons do not. There is nothing mysterious about the emergence of a property at some level, despite its complete absence at lower levels. Such emergence is ubiquitous.

The emergence of new properties in complex systems does not, however, mean that the whole is irreducible. While emergence is often defined as the appearance of wholes having irreducible novel properties, novelty and irreducibility do not need to be linked. Indeed, Anderson (1972) sees emergence and reduction as two sides of one coin, the difference being that emergence is what we see when we look up from the lower level to the higher one, while reduction is what we see when we look down from the higher level to the lower one. Emergence does mean that a whole is irreducible to its parts, but it does not require that the whole is irreducible to parts plus relations between parts, plus relations of whole and parts to the context in which they’re embedded.

The distinction between novel properties that are or are not reducible is the difference between “strong” and “weak” emergence. Nagel acknowledges this difference when he speaks of “harmless” (weak) emergence, but seems to assume that an explanation of mind via emergence requires the harmful (strong) variety. It is not clear why this should be so. Even strong emergence should not be objectionable if one can explain why reduction of emergent properties to lower level description is impossible. For example, it might be impossible because of mathematical undecidability or computational complexity or spontaneous symmetry breaking, or for other reasons. The often vacuous notion of “supervenience,” where novel upper level phenomena are said to depend upon but are still not reducible to lower level phenomena, might in this way be given some content.

**Life and mind are emergent but not purely materialist**

There is at least one reason to believe that an explanation of subjective experience is possible, even though we don’t have such an explanation now and don’t know where to look. The reason is the precedent of the scientific explanation of life. If, provisionally, we hypothesize that life emerged from matter and mind emerged from life, and consider only the constitutive question and not also the historical one, we note that both instances of emergence were once mysterious, but now only the mystery of mind remains. The solution of the first mystery suggests that solution of the second mystery is conceivable.

Although our understanding of life does not itself yield an understanding of mind, it calls attention to the fact that life is not a purely materialist phenomenon; the same might be true for mind. We understand life as a phenomenon in which organizational separation from an environment allows self-construction and repair (“autopoiesis”), using imported matter and energy processed by metabolism that is specified by internal information,
where this information also governs boundary creation, reproduction, and adaptation to
the environment, and where populations of such systems undergo evolutionary change.
These multiple aspects of life are fairly well understood, and few if any scientists would
say that we need a reductive theory that attributes proto-life features to the constituents of
living systems, to the constituents of these constituents, and so on. Even if one allows,
for the sake of argument, that some constituents (or their precursors, e.g., ribozymes)

evitably exhibit proto-life, one could not make this attribution for constituents even one level
down, since nucleotides or amino acids are not living in any sense of the word. Biology

offers no support to a reductive pan-life position. Life does not go all the way down.
Those who think otherwise typically confuse energy or dynamics with life. If life is

emergent, it seems reasonable to expect that mind is as well. It is hard to imagine that
mind has a reductive explanation and goes all the way down, while life does not.

But Nagel does not accept life as emergent, because a constitutive explanation of life
does not answer the historical question of how life appeared. Nagel takes the absence of
an account of origins as undermining not only our historical understanding but also the
constitutive explanation. This is unreasonable. Origins are always difficult to explain.
The origins of the universe or its galaxies or the solar system or its planets are obscure,
but this does not completely prevent understanding of diachronic change, nor does it put
in doubt our knowledge of the synchronic structure and function of these entities. The
absence of an account of origins does not undermine historical explanation, and the
incompleteness of historical explanation does not invalidate constitutive explanation. (It
should be mentioned that Nagel is also unimpressed by evolutionary knowledge about
life after it originated, so it is not only the absence of an account of origins that he thinks
flaws our historical explanation.)

The solution available to the constitutive problem of life suggests that the constitutive
problem of mind might be solvable, and that the solution may not be purely materialist,
since the constitutive conception of life is not purely materialist. Nagel’s characterization
of neo-Darwinism as materialistic is incorrect, however conventional it is. Life is based
not merely on a particular materiality, but on certain forms and processes. Autopoiesis,
replication, adaptation, and evolution can all be conceptualized without reference to any
specific materiality. This is the core insight of the field of “artificial life” (Langton
1997). In Bunge’s (1973) terms, the properties defining life are “stuff-free.” They are

capable in principle of realization via different forms of matter, although the instantiation
of these properties on Earth in fact requires organic molecules. Even if life everywhere
(assuming it also exists elsewhere) requires this materiality, it would still be a mistake to
regard the phenomenon of life as purely materialist, since these signature properties are
formal, not material. Autonomy depends on the self-reference of autopoiesis (Maturana
& Varela 1980), and self-reference is not a material property. The principles of evolution
are formal, not material, as the genetic algorithm illustrates. Understanding life requires
ideas of information to explain genetic determination, positive and negative feedback to
explain biological control, and utility and fitness to explain agency and evolution. Unlike
matter and energy, information and utility have no physical units. A theory that invokes
properties that have no physical units is not purely materialist. Of course, information
and utility must be instantiated, so their role in life, while not purely materialist, is not
idealist either (certainly not idealist in the senses of Table 1). Matter and form are joined
together, as Aristotle held. But Plato was also right, not necessarily in his idealism, but because his “vertical” separation of matter and form captures the fact that biological information occurs in various degrees of refinement, distillation, and concentration.

But abandonment of reductionist materialism, while necessary, is still not sufficient to explain subjective experience. There are numerous conceptions of mind that are formal, as opposed to material. Dennett (1991) joins a functional account of mind to eliminative materialism; Tononi (2008) offers a pan-psychic view based in information theory that echoes a proposal by Chalmers (1996) and is in the spirit of Nagel’s reductive neutral monism; and there are other functional, informational, or computational conceptions of mind. But the mystery of subjective experience remains unsolved by these non-material approaches.

**Mind is one of several instances of informational emergence**

While the formal approach to life doesn’t solve the mystery of mind, it does suggest a way to retain a modified reductive viewpoint. One can regard subjective experience as an emergent novelty, but see this emergence as continuous with other informational aspects of living and even non-living matter. This continuity could be described in terms of a hierarchy of levels of information, as shown in Table 2. This hierarchy is introduced to convey two ideas: (1) that types of information can be defined as going from general to specialized, encompassing the domains of matter, life, and mind, and (2) that subjective experience is a specialized process that occurs only at the highest levels of the hierarchy.

**Table 2. Levels of information**

The list of system types is borrowed (with slight modification) from Boulding (1956)

<table>
<thead>
<tr>
<th>SYSTEM TYPE</th>
<th>INFORMATION</th>
<th>PHENOMENON</th>
</tr>
</thead>
<tbody>
<tr>
<td>(vii) humans</td>
<td>neurological</td>
<td>self-awareness</td>
</tr>
<tr>
<td>(vi) animals</td>
<td>neurological</td>
<td>experience</td>
</tr>
<tr>
<td>(v) genetic-societal level</td>
<td>meta-genetic</td>
<td>differentiation</td>
</tr>
<tr>
<td>(iv) cells</td>
<td>genetic</td>
<td>reproduction</td>
</tr>
<tr>
<td>(iii) autopoietic systems</td>
<td>network</td>
<td>production</td>
</tr>
<tr>
<td>(ii) control mechanisms</td>
<td>control</td>
<td>regulation</td>
</tr>
<tr>
<td>(i2) clockworks (dynamic systems)</td>
<td>algorithmic</td>
<td>organization</td>
</tr>
<tr>
<td>(i1) frameworks (static systems)</td>
<td>form</td>
<td>organization</td>
</tr>
</tbody>
</table>

In this hierarchy, each level is both similar to and different from adjacent levels. Emergent neutral monism is reflected in the differences, and reductive neutral monism is reflected in the similarities. Subjective experience emerges at level (vi), and level (vii) represents its fuller development in consciousness. These two levels are genuinely new, but they are also a variation on informational phenomena occurring at lower levels. The reductive view is correct as applied to pan-information, but not to pan-psychism. Information goes all the way down, but psyche, i.e., subjective experience, does not. Psyche emerges.

The lowest level of the table shows information in its most general manifestation (i1), namely “form,” measured in information theory by Shannon entropy (Shannon & Weaver 1949), and the dynamic version of information (i2) developed in algorithmic information
theory of Kolmogorov (1965) and Chaitin (1975). Level (ii), information occurs in control systems, where it is concentrated and explicit. At level (iii), self-constructing but non-living systems (e.g., flames, eddies, tornadoes, which are autopoietic in a narrow sense) depend on a network of interactions that construct and maintain the internal order of the system. At these three levels, information is analog. In cells, genetic information at level (iv) governs metabolic autopoiesis and enables reproduction and adaptation. Here for the first time, information is digital, and here reproduction produces populations that have history, i.e., undergo evolution, which, contra Varela and Maturana, is an essential aspect of life. In the dynamics of cellular metabolism, enzymes are (analog) informational catalysts at level (i) that are specified by the (digital) genetic information at level (iv). Cells also include feedback control mechanisms, assigned here to level (ii). In multi-cellular plants and simple animals, there is the meta-level control (v) of genetic information by chemical messengers. Finally, at some point in the animal kingdom, one has complex neural systems (vi), that allow the emergence of subjective experience, and in higher animals (vii), self-awareness. Table 2 is just a crude sketch. One might wish to interpose between (vi) and (vii) a level for simple neurological systems that are sub-experiential, or add a top level (viii) that includes socio-cultural (language) phenomena.

This view of mind, which combines emergent neutral monism with a modified reductive view, might be called “continuous emergence,” where “continuous” is intended to have two meanings. It points to emergent phenomena that resemble lower level informational phenomena, and it suggests the repeated occurrence of emergence resulting in a hierarchy of levels, each both similar to and different from the level below.

Identifying the sense in which the reductive view is correct, namely the informational character of subjective experience, does not actually explain the emergence of level (vi) from level (v). It merely acknowledges that in emergence there is also continuity. It does argue, however, against taking lower level explanations as fully adequate for higher level phenomena, i.e., it rejects “systems-theoretic reductionism,” which is not an oxymoron. Reductionism of any kind explains phenomena in terms of something “fundamental.” In materialist reductionism, the fundamental is what is materially elemental; in systems-theoretic reductionism, the fundamental is what is most general, i.e., level (i) in Table 2. Systems-theoretic reductionism is illustrated by Wolfram’s (2002) argument that virtually everything dynamic can be modeled by cellular automata; among mind-matter explanations, it is exemplified by Tononi’s (2008) proposal that consciousness is “integrative” (holistic) information processing. But this argument against the sufficiency of reductionist explanations does not deny their importance or necessity.

The levels of Table 2 do not reflect a simple historical account, being more like a list of levels of “(informational) being,” but in so far as the instantiation of information in matter, life, and mind appeared sequentially in history, the table raises the question of teleology: is there, as Teilhard de Chardin (1959) thought, a tendency in cosmological evolution towards the distillation, refinement, and concentration of information? Nagel also asks us to consider the possibility of teleological laws that supplement causal laws, though it is unclear why additional laws are need if one posits an ample version of the Anthropic Principle. But these questions are beyond the scope of this paper to engage.
Summary
Continuous emergence advocates an neutral monism that takes subjective experience to be absent in simpler systems. The emergence of such experience, however, is one of several emergent informational phenomena, that range over the domains of matter, life, and mind. In Nagel’s terms, a neutral monism based in emergence can be augmented by a reductive pan-information, but not pan-psyhic, view. But this dual emergent-reductive framework does not solve the mystery of subjective experience, as it does not answer either the constitutive question of mind or the historical one.

References


