How to construct consensus models to (maybe) make sense of the mind-body problem

Martin Korth*

IVV NWZ, WWU Münster, Wilhelm-Klemm-Str. 10, 48149 Münster (Germany)

* Corresponding author email: dgd@uni-muenster.de

v2 2025/05/13; v1, 2025/02/10

A recent article by Kuhn[1] showcases the plethora of proposed solutions for the mind-body problem as it is understood in current 'consciousness science'. Perusing this article, philosophers will likely find it surprising to see the inclusion of for instance Indian idealism and Buddhist thought, but neither German, nor British or US idealists, which seems especially unbalanced when instead of them theories like Kastrup's analytical idealism (Hegel for physicists?) or Hoffmann's interface theory (Kant for psychologists?) are included. The listings of dualist, panpsychist and further theories seems equally eclectic, with for instance Descartes no more than a side-note in a number of entries, but several listings of religious ideas of body and soul. A possible problem with such omissions is of course that not taking philosophical history more seriously comes with the peril of repeating (now) avoidable errors. Nevertheless, judging from the disclaimers, Kuhn is well aware of this and most likely sees his list as more of a 'sociology of consciousness science', i.e. a list of people and ideas that are currently circulating in this field.

Four ways to approach the problem ...

It can in any case be taken from Kuhn's monumental work effort, that it is often assumed that a solution to the mind-body problem would be either (in the wider sense) materialistic (including non-reductive and quantum theories), dualistic, panpsychistic (including – nontrivial, i.e. not simply materialist or idealist – monist theories) or idealistic. The core problems faced are then those of emergence (of consciousness), interaction (between brain and mind), de/combination (of 'mind dust' or a unity into proper subjects) or emanation, where the emanation problem is the problem of explaining in detail(!) how the physical world and our scientific theories arise from a given ide-

alistic setup. To me it seems more likely that all these views have a point, but that none manages to cover the whole picture – and that it therefore might be possible to integrate these views into 'consensus models', if we do not look at the proposed solutions, but the initial intentions and/or intuitions for setting them up (much like as if we would mediate a social conflict).

... and what to learn from them

While materialists (in the wider sense) propose that the phenomenon of consciousness has to be explained on the basis of (current) physics, the actual intention of many materialists will be to stay in line with scientific inquiry. And while dualists propose that mind and matter belong to different spheres, their actual intention will in many cases be to explain the perceived mind/matter gap of human experience. Furthermore, whereas panpsychists propose a more uniform integration of mind and matter, often their intention will be to explain the seamless development of conscious beings from what seems to initially be no more than matter. Finally, idealists propose that reality is reducable to mental phenomena, mainly to make space for phenomenological experience, universal abstract entities and proper subjectivity.

Four solutions available ...

Ironically, the different approaches are even able to deliver 'solutions' for the core problems of the other ones: The strong subject of idealism could be understood as a solution to the combination problem of panpsychism, i.e. a binding focus for the disparate non-material properties that make up a mind. The seamless development in panpsychism could in turn deliver an answer to the interaction problem of dualism, in the sense that the observed gap could be explained by a diverging development from initially uniform building blocks. And the resulting dualism could then explain why reductionist physics is so successful in the material world, but has to be extended for the non-material one. Finally, re-interpreted physicalist theories would be the solution to the emanation problem of idealism.

... and four traps to avoid

Nevertheless, all approaches also carry self-inflicted wounds from their history with them:

Physicalists often confine themselves to current physics, although physics was able to suc-

cessfully re-invent itself several times now. So why not test hypothetical models with more than the currently understood 'material' properties (like for instance charge or spin), if the additional properties (like for instance colors) can afterwards be supported by proper scientific inquiry? Properties that are currently understood to be 'non-physical' do not have to be a fundamental problem for an evolving physics.

Dualists in turn often overlook that a non-fundamental (developmental) mind/matter gap would have an equally strong explanatory power as a simply given one. The therefore required revision of also our theories for the material world then offers more opportunities than a simple addition of a mental sphere to known physics when trying to solve the interaction problem. The interaction problem does thus not have to be a fundamental hindrance for ('effective') dualists.

Panpsychists often limit the mental to (especially phenomenological) content, leaving out proper (individual) subjectivity; 'atomic' panpsychists then encounter the same problems as dualists (how does a mind form?), 'holistic' panpsychists the same as traditional idealists (how do separate subjects arise?). But the combination problem vanishes if some basic form of proper subjectivity is also taken as a building block.

Traditional idealists finally often resort to a quasi-religious idea of unity, which is not directly supported even by their own standards, and which lets them encounter the then self-inflicted 'decombination problem' of holistic panpsychism and makes the emanation problem (of how to get from there to physics, evolution and neurobiology) overly hard, as everything(!) specific has to be understood as some kind of a decoupled mode of a unified subjectivity field or world soul. But there is no reason why 'atomic idealists' could not start from objectively existing properties and very basic subjects to avoid this and thus make the emanation problem much easier.

Constructing 'consensus models'

To build upon the above we can therefore outline 'rules' for the construction of 'consensus models' based not on the proposed solutions, but on the intentions behind the different approaches:

From physicalism we can take it that consensus models will have to closely integrate science, i.e. explain (in detail!) why we arrive at the scientific theories that we have.

From dualism we can take it that consensus models will have to explain (again in all detail!) why we experience the world dualistic (idealist theories for instance usually don't do this).

From panpsychism we can take it that consensus models will have to explain how life and consciousness gradually develops from simple building blocks (physicalists explanations for this are often challenged when it comes to consciousness).

And from idealism we can take it that consensus models will have to explain why we experience ourselves as a subject, as well as why we find a 'third world' of seemingly subject-independent abstract entities.

A successful consensus model could then be called all of the above, depending on what parts one wants to emphasize.

The way forward

A first example for such a model, under the label of an 'pragmatic' or 'scientifically tenable' idealism is Model A, or 'A-world':[2] Starting from an atomistic idealistic setup of (essentially monist) properties and a population of extremely simple 'core subjects' as building blocks, i.e. an effective panpsychism, the model proposes a divergent early development towards a mind matter gap, i.e. an effective dualism, with a material world in line with current physics, i.e. an effective scientific realism, and then a gradual physical, biological and then cultural evolution of human thought.

A mind is then the composite of a 'world map' of abstract entities with a core subject, with the abstract entities as composites of non-material properties as 'atoms of thought', and anchored to the physical properties of a brain via its subconsciousness: 'In this model, a person is the totality of body and mind. The mind is the totality of the core subject and the structured bundle of universals that the core subject can perceive and manipulate. The structured bundle is the representation of the person's world, their 'world map' (not only to be understood spatially), through which the core subject interacts with the world. World map and core subject can only be meaningfully understood in combination as a (self-)conscious mind. A 'mapless' subject is without any conscious content, a 'subjectless' world map is only an abstract object. Parts of the world map of the core subject are also part of the world maps of sub-subjects, which make up the person's subconsciousness and which allow interaction between the core subject and the body: Parts of the world maps of the subsubjects are in turn part of bundles with physical properties that belong to the structure of the person's brain, whereby the respective sub-subjects can manipulate both the mental and physical properties of their bundles on an equal footing, albeit according to different, evolutionarily learned

rule sets. The brain as part of the person's body thus functions as an anchor for the non-material mind of the core subject. However, the person's body can only be understood as a hybrid entity; it receives its embedding in the causal world via the countless bundles of physical properties of which it consists of (organs, molecular structures, but ultimately elementary particles), but its unity as an entity only in the connection of these structures to structured bundles of universals in the mind of the core subject, which is only aware of its body in this form. Every perception or action thus builds a bridge between causal network and qualitative representation, which, as we shall see, is causal in nature, but whose evolutionary emergence must be understood as a creative act.'[2]

The way forward is to work out such hypothetical consensus models in such detail, that they become available for scientific inquiry, in the form of new and decisive experiments to be tested (thus hopefully helping to move beyond the endless iteration of old arguments and shuffling between known positions).

References

[1] R. L. Kuhn, *A landscape of consciousness: Toward a taxonomy of explanations and implications*, Progress in Biophysics and Molecular Biology, 2024, 190, 28-169. https://www.sciencedirect.com/science/article/pii/S0079610723001128

[2] M. Korth, Information, Intelligence and Idealism, 2025. https://philpapers.org/rec/KORIIA-4