More hydra, than human? Moral considerability of human brain organoids based on neural architecture

J. Lomax Boyd
Berman Institute of Bioethics, Johns Hopkins University, Baltimore, MD, USA
Email: lomax.boyd@gmail.com; Twitter: @lomaxboyd

1. Background

- Human brain organoids (HBOs) are novel entities grown from stem cells with uncharacterized cognitive potential to exhibit ‘intelligent’ features.
- The capacity for having intrinsic interests, or conscious experiences with positive or negative valence, are considered sufficient for moral status [4].
- Epistemological criteria for whether HBOs possess these morally-relevant capacities remain limited and contested.
- One approach—similarity of HBOs to human neuroanatomy—has notable scientific, technical, and philosophical limitations when applied to organoid-based entities [3].

2. Central Questions

- How should we conceptualize the cognitive potential of HBOs given the variability and novelty of their information processing architecture?
- What are the moral implications of different neuronal architectures?

Figure 1: HBOs equipped with multielectrode arrays (MEAs), which provide and record electrical stimulation used to embody organoids within virtual environments, can be used to measure functional connectivity [5].
3. Main Results

3.1 Major transitions in cognitive evolution can be used to benchmark the cognitive potential of novel biological entities

- Barron’s [6] theory of transitions in cognitive evolution explicitly recognize the role of neural diversity and phenotypic potential underlying organizational principles of the brain
- Theories of information processing provide a principled approach of relating (brain) structure to function (cognition)
- Comparison of HBOs functional connectivity to Barron’s five computational architectures can provide a framework for inferring the cognitive potential of novel biological agents.
- How should computational architecture inform moral considerability?

Adapted from Barron et al., 2019
3. Main Results
3.2 Moral consideration of HBOs depend on computational architecture and functional capacity

- Functional capacities, such as memory, scale at different rates depending on the underlying geometry of neuronal organization [2]
- Functional connectomes (or control flows) of first generation (FG) brain organoids are likely equivalent to decentralized architectures of the hydra, which are not known to support morally-relevant dimensions of cognition.
### 3.2 Main Results

3.3 Pathway toward experimental assessment of criteria underlying moral consideration [1]

<table>
<thead>
<tr>
<th>Computational Architecture</th>
<th>Cognitive Motif</th>
<th>Dimension of Cognition</th>
<th>Moral Category</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Example</strong>: Recurrent</td>
<td><strong>Example</strong>: Temporal error prediction</td>
<td><strong>Example</strong>: Evaluative richness</td>
<td><strong>Example</strong>: Evaluative stance</td>
</tr>
<tr>
<td>Question: Does the neural network of the organoid exhibit feedforward and feedback control over information flow?</td>
<td>Question: To what extent does the organoids neural network support error prediction?</td>
<td>Question: Does the entity exhibit evaluative decision-making?</td>
<td>Question: What moral status is afforded sentient beings with similar information flows and evaluative decision-making?</td>
</tr>
</tbody>
</table>
4. Conclusions

4.1. Computational architecture provides a novel framework for discovering epistemological criteria for moral status in novel biological entities, like human brain organoids

4.2. There remains notable uncertainty regarding the ontology of cognitive capacities, including those most relevant to moral status

6. Funding
This work was supported by a grants to Lomax Boyd from the Kavli Foundation and the Johns Hopkins Whiting School of Engineering and Applied Physical Laboratory (SURPASS)

5. References


