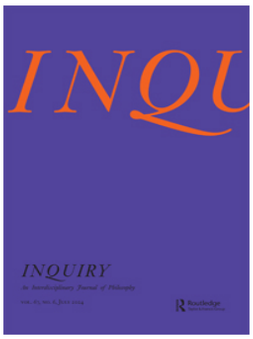

Routledge
Taylor & Francis Group

Inquiry
An Interdisciplinary Journal of Philosophy

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/sinq20

Counting subjects

Joseph Gottlieb & Bob Fischer

To cite this article: Joseph Gottlieb & Bob Fischer (02 Jul 2024): Counting subjects, Inquiry, DOI: 10.1080/0020174X.2024.2370016

To link to this article: https://doi.org/10.1080/0020174X.2024.2370016

Published online: 02 Jul 2024.

INQUIRY
https://doi.org/10.1080/0020174X.2024.2370016

Routledge
Taylor & Francis Group

Counting subjects

Joseph Gottlieb [a] and Bob Fischer [b]

[a]Department of Philosophy, Texas Tech University, Lubbock, USA; [b]Department of Philosophy, Texas State University, San Marcos, USA

ABSTRACT

We normally assume that there's just one conscious individual per animal. Some question this, suggesting that there may be nonhuman taxonomic groups whose normal, adult members house more than one conscious subject. Call this the *multitudes view* ('MV)'. Our aim is methodological: we hope to understand *how* we might assess whether MV is true. To that end, we distinguish two strategies for counting conscious subjects: the duplication strategy and the mind-first strategy. We use human split-brain patients and octopuses to illustrate the prospects and limitations of each approach, as some have argued that both sets of individuals contain multitudes. We'll contend that while the duplication strategy is applicable to split-brain patients, it isn't apt for octopuses. The mind-first strategy fairs better here, but there's a catch: it requires, in a way the duplication strategy does not, numerous contentious assumptions about the nature of mentality, the individuation of minds, and consciousness to simply get off the ground. These difficulties clarify just what it would take to establish MV.

ARTICLE HISTORY Received 11 March 2024; Accepted 15 June 2024

KEYWORDS Consciousness; split-brains; octopuses; phenomenal unity; animal minds; nonconscious mentality

1. The multitudes view

We normally assume that there's just one conscious individual per organism. When we point to a human, we point to precisely one conscious subject – one conscious individual – and likewise for pigs, platypuses, and poodles. However, some have raised questions about the universality of this principle, suggesting that there may be taxonomic groups whose normal, adult members house more than one conscious subject (equivalently, 'subject of experience'). This is important, both theoretically and ethically. Theoretically, it would allow us to make some inferences about the evolution of different forms of consciousness, potentially

CONTACT Joseph Gottlieb joseph.gottlieb@ttu.edu Department of Philosophy, Box 43092, Texas Tech University, Lubbock, TX 79409-3092, USA

identifying selection pressures that give rise to unified and disunified subjects within organisms (Veit 2023). Ethically, it recalls *Bentham's dictum*. One can agree that 'each [should] count for one and none more than one' (Mill 1861/1969, 257) – one can agree that what matters in moral deliberations is not *whose* interests we are considering, but the interests themselves – yet still wonder how to count in the more fundamental 'enumerating each one' sense. Every being deserving of moral consideration may count for exactly the same (one and none more), but in some cases it may be wrong to count each whole animal *for one*. They may instead count for many, precisely because, containing more than one conscious subject, they house more of what (fundamentally) counts.[1]

One way to motivate the idea that some contain multitudes begins with humans – so-called *split-brain patients* – who have had their corpus callosum severed. When different stimuli are presented to the left and right portions of their visual fields, these individuals can report what's visible on the right side and draw what's visible to the left side, but not vice versa. So, we might ask whether the *experiences* of the stimuli are disunified too – or, more radically, whether there are two *subjects* of experience housed in one skull.[2] Birch, Schnell, and Clayton (2020, 793) extend this line of thought to nonhumans:

> Birds … are natural split-brains. They have no structure akin to the corpus callosum connecting the two hemispheres of the dorsal pallium, which is homologous to the cortex in mammals. Could every bird be a pair of conscious subjects, intimately cooperating with each other? A similar debate arises with respect to the cerebral ganglia and brachial plexus (a nerve ring around the top of the arms) of the octopus. These structures are connected, but they have some degree of functional autonomy from each other. Could an octopus have two, or even nine, conscious perspectives on the world?

[1]One possibility that we'll encounter later is that an animal houses multiple *minded* subjects that *lack* consciousness. This is not the view of interest. But one might wonder why it isn't, even with the ethical angle; after all, some doubt that consciousness has ethical salience (e.g. Carruthers 2005; Levy 2024). Perhaps having a mind is enough for welfare subjecthood and moral considerability. Although not a popular view (see, e.g. Kriegel [2019] for an overview), we think there is something to this idea. However, even among those who take this position, it's often granted that consciousness nonetheless *enhances* the value of certain welfare states, making it the case that those with the capacity for conscious experience have more at stake in survival (e.g. Bradford 2023). Thus, it may not only be that some animals don't count for one, but it may also be that the (multiple) ones they count for count for a lot!

[2]Schechter (2018; 2022) gives the most extensive and compelling defence of an affirmative answer to this question. Unsurprisingly though, there is still much controversy. Indeed, 'single subject views' disagree amongst themselves – over the hemisphere with which the (single) subject of experience is associated, what to say about the remaining hemisphere, and so on (e.g. Tye 2003; Bayne and Chalmers 2003; Bayne 2010). The good news is that, for present purposes, Schechter does not need to be *right* that split-brain patients contain two subjects. Our contention is that even if human split-brain patients do *not* contain multitudes, getting to that answer will, along the way, have invariably taught us quite a bit from a strategic perspective when it comes to MV.

Let the *multitudes view* ('MV') be the thesis that there are nonhuman taxonomic groups whose normal, adult members house more than one phenomenally conscious subjects at a time, as part of their default condition. This paper explores the prospects for building a case for MV.

MV requires some unpacking. By 'house' we mean whatever relation obtains between the whole animal – a member of the target taxonomic group in question – and conscious subjects, be it parthood, constitution, composition, or something else. So stated, MV is neutral on the nature of this relationship. By 'default condition' we mean something like *in the wild*. This is to rule out, for instance, possible cases where multiple subjects arise in experimental contexts only to disappear when outside those contexts.

By 'phenomenally conscious subject' we mean an entity that is non-derivatively phenomenally conscious. Phenomenally conscious subjects are such that there is something it is like *to be* them. Mental states, like smelling cigar smoke or hearing a piano, are like something *for* them. When a subject S is phenomenally conscious *of* cigar smoke – has a conscious experience *of* cigar smoke – this is because S is in an olfactory mental state that represents the smoke and is itself phenomenally conscious.[3] All of that is standard fare. Let A be some animal and E1 and E2 be two distinct conscious experiences – an olfactory experience of cigar smoke and an auditory experience of piano music, say. MV's distinctive claim is this: A derivatively has E1 and E2 in virtue of housing things (multiple subjects) that non-derivatively have E1 and E2 (Schechter 2018, 81–82).[4]

While some highly speculative metaphysical considerations might favour MV (see, e.g. Crummett 2022; Gottlieb and Fischer 2024), we're interested in MV's empirical motivations. The split-brain condition in

[3]Going forward, the 'phenomenal' qualifier will be dropped, save for emphasis.

[4]The idea that things can have properties derivatively is familiar from the literature on constitution-without-identity. Here is an example from Baker (2000, 46ff). Suppose we have Michelangelo's *David* and 'Hunk', the hunk of marble that constitutes it. *David* and Hunk share properties. Both *David* and Hunk are over 6 feet tall. Yet David is over 6 feet tall only because it is constituted by something (Hunk) that is over 6 feet tall. Hunk's being over 6 feet tall holds independently of its coming to constitute David. In that way, Hunk has this property non-derivatively. The key point here is the direction of explanation, not whether A and the multiple subjects it houses stand in the same relations as David and Hunk; if A has an experience of cigar smoke it will be *in virtue of* housing something that experiences cigar smoke.; Given that MV defines subjects as being non-derivatively conscious, how should we interpret attributions of consciousness to MV candidates? Let a *n-attribution* of consciousness to some animal *x* be an attribution of consciousness to *x* that is *neutral* on whether *x* qua whole animal is non-derivatively conscious or whether *x* is derivatively conscious in virtue of having as part, constituting, or otherwise housing some entity *x'* that is non-derivatively conscious. When we make whole animal attributions (e.g. 'split-brain patients are conscious' or 'octopuses are conscious') without qualification, we are only making n-attributions of consciousness.

humans has arguably generated the most interest in the possibility of a single human housing multiple conscious subjects. Yet even if humans (split-brain or otherwise) house just one conscious subject, this doesn't preclude non-human animals from housing more than one (Bayne 2008, 300). Indeed, the lack of integration and lateralisation exemplified in the split-brain condition is often the *norm* in other animals, including cetaceans (Roth and Dicke 2017, 142), birds (Vallortigara 2000), and lizards and fish (Godfrey-Smith 2020, 239). And because human consciousness is better studied and understood than non-human consciousness, the split-brain condition can help us better understand *how* to ascertain whether MV is true of various nonhumans.[5]

In this way, the central aim of this paper is comparative and methodological. MV invites *the counting question*: how should we count conscious subjects? If MV is even possibly true, it can't be that we count conscious subjects by counting (perhaps some subset of) whole animals. So, another approach is needed.

We consider two. The first – what we'll call *the duplication strategy* – can be found (implicitly) in Schechter's (2018) investigation of split-brains. In general, to establish that an animal houses multiple subjects of conscious experience, we can, with less artificiality than it may first appear, factorise our aims into two parts: establishing that there are multiple relevant conscious experiences and establishing that these conscious experiences are had by numerically distinct subjects. As we'll see, a primary benefit of the duplication strategy is that it completes the former task in a comparatively brisk way, and it does so with minimal assumptions about the nature or function of consciousness. The downside, however, is that the duplication strategy is not apt for nonhuman MV candidates. In particular, it's not apt for what is perhaps the most interesting MV candidate: the octopus.[6]

A second approach – what we'll call *the mind-first strategy* – can be applied to the octopus but requires a stand both on the nature of mindedness and the conditions under which a mental state is conscious as opposed to nonconscious. While there are glimpses of the mind-first approach in Carls-Diamante's (2017, 2019, 2022) pioneering case that octopuses *do* contain multitudes, we'll see that failing to fully separate

[5]Pace Nagel (1971), we assume that there will always be some non-stipulative, objective answer as to how many conscious subjects an animal houses. There are facts of the matter.

[6]MV is not the only topic of note when it comes to the octopus; the octopus's array of remarkable cognitive abilities has also attracted attention. See, e.g. Birch, Schnell, and Clayton (2020); Carls-Diamante (2017, 2019, 2022); Godfrey-Smith (2016, 2020, 2021). For a critical take on the supposed disunity of the octopus's nervous system, see Mather (2021).

the duplication and mind-first strategy, along with failing to appreciate the theoretical burdens of the latter, render her argument unpersuasive. While the duplication strategy *cannot* show that octopuses contain multitudes, and our current evidence says that the mind-first strategy *doesn't*, the main methodological headline is that the mind-first strategy at least *could*.

2. The counting question

Both the duplication and the mind-first strategies consist of two core steps: a *consciousness step* and an *individuation step*. They differ in terms of what's involved in each step, and critically, the order in which each step is taken. On neither strategy though do we need to be told what conscious subjects *fundamentally are*, and indeed we do not interpret them that way. For these strategies to be useful, they only need to provide reliable proxies for subjects such that we can count them. One can do so while remaining to at least some extent neutral on the deeper metaphysical question. After all, if MV were false, and we *could* count subjects by counting animals, this would not mean that subjects *are* animals. We begin with the duplication strategy (§ 2.1), before turning to the mind-first strategy (§ 2.2).

2.1. The duplication strategy

We've already identified the basic motivation for investigating whether the case for multiple subjects in human split-brain patients carries over to nonhumans: namely, there are interesting similarities between split-brain patients and certain nonhuman animals, and some theorists have taken these similarities as a central basis for exploring whether these nonhumans house multitudes (e.g. Birch, Schnell, and Clayton 2020; Godfrey-Smith 2021).[7] The duplication strategy permits what is perhaps the most powerful case for duality in split-brain patients. Thus, it is worthwhile to investigate the duplication strategy in particular, to see not just why it ultimately cannot be extended to nonhumans of interest, but also how its shortcomings point to a more promising way forward.

That case begins with the notion of *phenomenal unity*. Carmichael walks into the Blue Note. Carmichael's senses are inundated with a

[7]While we think that Carls-Diamante can be profitably understood as pursuing the mind-first strategy (§ 2.2), this is not an entirely consistent assessment. See § 3.2 for more.

range of stimuli. Carmichael *smells* cigar smoke; Carmichael has an *olfactory experience*. Carmichael *hears* a middle C on the trumpet; Carmichael has an *auditory experience*. But he also *sees* many things. Carmichael sees – and hears too – the man playing the piano. If Carmichael were to describe what it is like to be at the Blue Note, he could list these experiences one-by-one and describe what they are like. Yet this would yield a phenomenologically impoverished description. There is something that it is like for him to undergo these experiences together at a time that is over and above what it would be like for Carmichael to undergo any of these experiences on their own. If he were just to point out that there is something it is like for him to smell the cigars, *and* hear the trumpet, *and* see the pianist, the description would remain phenomenologically impoverished. The proper description is not conjunctive, but one that appeals to a *unity* in our mental lives. This unity need not itself be an *object* of experience – something like the cigar smoke, or the pianist – but it makes a phenomenological difference. Carmichael's experiences are unified insofar as they are subsumed by a single conscious state. And this subsumption yields a distinct conjoint phenomenology.

Carmichael's experiences are *phenomenally unified* ('p-unified'). Let two conscious mental states E1 and E2 be p-unified at some time *t* if and only if there is something it is like to undergo E1 and E2 *together* at *t*. What it is like to undergo E1 and E2 together at *t* does not reduce to what it is like to undergo E1 at *t* and what it is like it is like to undergo E2 at *t*.[8]

Carmichael's experiences are also *access unified* ('a-unified'). This is an extension of access consciousness ('a-consciousness'), where a mental state is a-conscious if its content is available for use in reasoning, reporting, and rational guidance of action (Block 1995). Carmichael can jointly report on, and jointly employ in the rational control of reasoning and behaviour, not just visual information about the cigars (brownness and cylindricality), but information about the cigar smoke, and the sound of the piano. A bit more formally, we can say that two conscious mental states E1 and E2 are a-unified at some time *t* if and only if the *conjunction* of their contents is available for verbal report, reasoning, and the deliberate control of behaviour at *t* (Bayne and Chalmers 2003). If E1 has content <The piano is black> and E2 has content <The cigar is brown>, then E1 and E2 are a-unified a *t* iff <<The piano is black> ∧ <The cigar is brown >> is available for report and control at *t*.

[8]This is p-unity 'by subsumption' (Bayne and Chalmers 2003). P-unity is symmetrical (if E1 is unified with E2, E2 is unified with E1) and reflexive (every experience is unified with itself).

Now consider three theses:

Subject Unity

E1 and E2 are subject unified ('s-unified') insofar as they are (non-derivatively) had by the same subject.

S-unity ⇔ P-unity

If E1 and E2 are (non-derivatively) had by a subject S at a time *t* iff E1 and E2 are p-unified at *t*.[9]

Ownership

All experiences are (non-derivatively) had by a subject.[10]

Suppose we have an animal MV candidate, *A*. Further suppose that A has at *t* – if only derivatively – conscious states E1 and E2. If E1 and E2 are *not* p-unified, then E1 and E2 are not non-derivatively had by the same subject. Since *A*, we'll assume, is a unitary animal, at most one of E1 and E2 would be non-derivatively had by *A*. Given **Ownership**, there must be some further subject, S, that is undergoing either E1 or E2, who is not identical with *A*, albeit perhaps a part of or otherwise housed by *A*. Another option is that *A* only derivatively has E1 and E2, and is composed or otherwise houses *two* subjects, S and S', one who non-derivatively has E1 and another who non-derivatively has E2. Either way, given **S-unity ⇔ P-unity**, we can use the absence of p-unity to get multiple subjects. So, if we can assess whether we have failures of p-unity, we can count conscious subjects.

Two questions now emerge. First, notice that the application of both **Ownership** and **S-unity ⇔ P-unity** requires that we first establish the presence of consciousness; without conscious experiences, there will be nothing to pin a subject to, and without conscious experiences, there is nothing for p-unity to be a relation between. So, to establish the presence of consciousness we need a fourth principle:

The duplication principle

If X is sufficient for conscious experience, and Y is an intrinsic duplicate of X, then Y is sufficient for conscious experience.

[9]See, e.g. Bayne and Chalmers (2003, 27), Dainton (2008), and Bayne (2010, 16).
[10]**Ownership** is meant to be *obviously* true. It's compatible with just about any view on the metaphysics of subjects or selves, even Humean bundle theories (e.g. O'Conaill 2024).

What can you do with the duplication principle? In the context of hemispherectomy patients – i.e. those who have had an entire cerebral hemisphere either removed or functionally incapacitated – the answer is, 'Quite a lot'.

Following Schechter (2018), we can say that, in addition to the right (RH) and left (LH) hemispheres of the brain, we have the *right hemispheric system* ('R') and the *left hemispheric system* ('L'). The hemispheric systems are not hemispheres; R is not just the right hemisphere ('RH'), and L is not just the left hemisphere ('LH'). Rather, Schechter (2018, 82) defines R as the whole animal without LH, while L is the whole animal without RH. As she puts it, 'R and L are physically distinct but not physically discrete proper parts of *A*' (Schechter 2018).

But now notice: R and L are intrinsically equivalent to hemispherectomy patients (Schechter 2018, 82). This is a powerful observation. In the present scenario, X is someone who has recently undergone a left or right hemispherectomy, and Y is R (or L). We *know* that hemispherectomy patients, be it derivatively or not, enjoy conscious experiences. So, by **the duplication principle**, R (or L) is sufficient for conscious experience.[11]

The expression 'is sufficient for conscious experience' is awkward but necessary. The duplication principle cannot on its own be used to establish that R and L are distinct conscious subjects. Here is Schechter (2018, 83):

> None of this shows that R and L actually are distinct [subjects], however. It may be that when both hemispheres are present and functioning, they inevitably interact sufficiently intimately so as to be brain structures of a single [subject]. The "how many minds?" question can only be answered by examining the actual nature of interhemispheric psychic interaction after split-brain surgery.

But the duplication principle is limited in another way: it can't even be used to establish that something *is* a subject. Suppose you are a soul theorist who thinks that all adult humans have a soul intrinsically. Then you could let X in the duplication principle be an adult human and agree that both X and Y are sufficient for conscious experience. This is not, however,

[11] While the externalist about consciousness (e.g. Dretske 1995; Tye 2009) might have qualms about the duplication principle, this worry can be dodged. For in the cases that matter, the relevant external conditions – like having an evolutionary history in which one's perceptual systems have the function of tracking things in the environment – aren't going to differ in any relevant way when comparing hemispherectomy patients with R or L. Another issue, one that's more of a quibble, is that we could have framed the duplication principle in terms of *capacities* for conscious experience. While this framing has its benefits, we set it aside here.

because X and Y *are* subjects of experience. Souls are, which are bought along whenever there is an adult human.

What the duplication principle *does* plausibly allow us to infer is that R and L are distinct 'loci' of consciousness – a neutral term for when something is the physical source of an experience – but we *cannot yet conclude that that something is a unique subject of that experience*. Thus, the consciousness step of the duplication strategy is simply the application of **the duplication principle**.

Second, **S-unity ⇔ P-unity** invites us to employ the absence of p-unity to infer multiple subjects. How though do we know whether two experiences are p-unified? This requires a fifth, and final, principle:

P-unity ⇔ A-unity

E1 and E2 are synchronically p-unified at *t* iff E1 and E2 are synchronically a-unified at *t*.

This principle tells us that if two experiences are *not* synchronically a-unified at *t*, then they are not p-unified; but if they *are* synchronically a-unified, then they *are* p-unified. Putting this together with **S-unity ⇔ P-unity**, we now have the individuation step.

For sake of completeness, here is how this can be put to work in the case of split brains. Let P be a particular split-brain patient. In P, there is the RH-set of experiences and the LH-set. The 'RH-set' (or 'LH-set') is the set of experiences whose members are the result of a suitably lateralised stimulus, such that information about that stimulus is passed only to a single hemisphere (the RH or LH). The justification that there is an RH-set and LH-set is **the duplication principle**.[12] P contains two hemispheric systems, R and L, both of which are intrinsic duplicates of something we

[12]Schechter doesn't hang her hat on just the hemispherectomy data and the duplication principle (Although she is not the only one to appeal to hemispherectomy data; see, e.g. Roelofs 2019, 204). She also appeals to independent empirical considerations concerning RH-based activities and whether they exhibit various behavioural markers of consciousness. (This is in defence of what Schechter calls the *bihemispheric consciousness claim*, which says that 'both hemispheres are the source or site of elements of experience' [2018, 27–28; 31–32]. Note that the duplication principle is also being used to establish the bihemispheric consciousness claim here; the source of experience in a hemispherectomy would presumably still be their remaining hemisphere.) Though not without merit, we do not recommend this approach if it is taken up *in lieu of* the duplication principle. For instance, Schechter mentions how a 'number of RH-controlled behaviours are of types taken to be decisive evidence of conscious experience in other contexts' (2018, 32). Outside of verbal reports – which is inapplicable when we consider nonhumans – it's not clear that there are *any* behavioural markers which are taken to be decisive evidence of conscious experience (see also §2.2). And although we will examine the so-called *theory-light* approach to consciousness later (§ 3.2), we doubt that the search for truly theory-invariant behavioural markers – behaviours which otherwise different theories of consciousness by-and-large agree can be proxies for consciousness – will be fruitful. This stands in contrast with the duplication principle, which is not to the same extent beholden to substantive theoretical assumptions (see fn. 11).

know is sufficient for consciousness (a hemispherectomy patient), which have RH and LH as proper parts, respectively.[13] Notice though there are no commitments yet about the *identity* of the subject of these experiences, only that they will have *a* subject (given **Ownership**).

Next: the individuation step. Roughly, Schechter's (2018, 24ff) argument goes like this. For any two experiences E1 and E2, E1 and E2 are part of the same 'experiential perspective' (EP's) if and only if E1 and E2 are synchronically p-unified. Yet E1 and E2 will be synchronically p-unified if and only if they are synchronically a-unified. This is **P-unity ⇔ A-unity**. And it turns out that members of the RH-set are synchronically a-unified with each other, and the members of the LH-set are synchronically a-unified with each other; as such, they are synchronically p-unified. The p-unified LH-set, and the p-unified RH-set comprise EP's. So, it follows that P's RH realises experiential perspective EP1 and P's LR realises experiential perspective EP2.

Schechter takes these observations to establish an *intrahemispheric unity claim*: that there are *at most* two perspectives, i.e. EP1 and EP2. However, the members of the RH-set are *not* synchronically a-unified, and so not synchronically p-unified, with the members of the LH-set. Thus, Schechter concludes, EP1 ≠ EP2. As such we also have *interhemispheric disunity*; there are *exactly* two perspectives (Schechter 2018, 29). Since, by **Ownership**, each EP is the perspective of some subject of experience, and each subject of experience non-derivatively has exactly one EP – no single subject can (non-derivatively) have two EP's in the *very same sense* in which each of us has *one* (Schechter 2018, 49) – it follows that P houses two subjects of experience, one associated with RH (R), and one associated with LH (L).

Never mind whether this is the correct interpretation of split-brains; never mind whether **S-unity ⇔ P-unity**; and never mind whether **P-unity ⇔ A-unity**.[14] What matters now is that we have a reasonable

[13]There is technically a leap here, since of course R and L have other proper parts – toes, kidneys, etc. – and we wouldn't thereby speak of toe-set of experiences or kidney-set of experiences. (Thanks to Elizabeth Schechter for this point.) But this shouldn't matter to what follows. If one likes, we can think of the 'RH-set' as referring to the R-set, and the 'LH-Set' as referring to the L-set, since, strictly speaking, **the duplication principle** tells us that R and L are sufficient for conscious experience.

[14]Both **P-unity ⇔ A-unity** and the weaker **P-unity ⇒ A-unity** are contentious, as they rely on controversial assumptions about the relationship between phenomenal consciousness (relevant to 'experiential perspectives') and access consciousness (Block 2007). Schechter (2018, 45) is clear, however, that split-brain patients are both p-unified intrahemispherically and p-disunified interhemispherically, and that our main line of evidence of this intrahemispherical a-unity and interhemispherical a-disunity, respectively. For an interpretation of split-brains where there is p-unity across the LH and RH but no a-unity, see Bayne and Chalmers (2003).

approach to MV, and the counting question it engenders. In sum, the strategy goes like this:

The Duplication Strategy:

1. **Step One** (the consciousness step): Employ **the duplication principle** to establish multiple loci of consciousness.

2. **Step Two** (the individuation step): Posit **P-unity** as an individuation relationship, letting p-unified EPs be proxies for subjects.

3. **Step Three**: Employing **P-unity** ⇔ **A-unity**, argue that p-unity fails to hold because a-unity fails to hold.

The question now is whether we can use the duplication strategy for a MV candidate like the octopus.

The answer is *no*. The impediment here is not that the octopus isn't conscious; it likely is (e.g. Mather 2008). Rather, the impediment is that we cannot get past Step One; the duplication principle is inapt. The idea is that the octopus houses up to nine subjects: one for the central brain, and one for each arm. But there are no intrinsic duplicates here. The arms are not an intrinsic duplicate of anything of which, we know, is sufficient for consciousness.

Of course, this doesn't show that the octopus fails to substantiate MV. It just shows that a different strategy is needed. We outline this new strategy next.

2.2. The mind-first strategy

The duplication strategy had a consciousness step, and an individuation step – *in that order*. The *mind-first strategy* shares these steps, but in the opposite order. This means that we can identify proxies for subjects, and count them, *prior to* knowing whether these subjects harbour conscious states.

Two points of clarification on this. First, the mind-first strategy counts subjects by counting minds. Some of these subjects may have conscious experiences like you and I, or perhaps a more limited range of experiences, but many may very well lack the capacity for conscious experience altogether. That's not to say, though, that conscious subjects – the entities of interest on MV – and what we might call *merely minded* subjects belong in fundamentally different ontological categories. The difference, per the mind-first strategy, is just a matter of what these things are subjects *of*.

Conscious states are mental states that have the property of being conscious (whatever that property ultimately is; the property of being in a global workspace, being HOT-targeted, etc.). Conscious subjects are subjects of conscious states; merely minded subjects are subjects of mental states that simply lack the property of being conscious.

Second, and relatedly, the mind-first strategy in this way presumes that there is nonconscious mentality. This is why it's hardly dispositive to point to the octopus arm's behaviour as evidence of MV. For even if this behaviour requires a *mental* explanation – which is itself not obvious – it doesn't follow that this explanation must be cashed out in terms of *conscious* mentality. We know – for quite general reasons – that mental states of arguably all kinds can be nonconscious.[15] Often what *looks* like consciously driven processes frequently aren't. That a conscious state and a behavioural pattern are correlated doesn't imply that this conscious state causes that behaviour. And even if a conscious state *does* cause some behaviour, that doesn't imply that it causes it *in virtue* of being conscious (Michel 2023), which means that behaviour could be realised by the unconscious counterpart of that conscious state.

Moreover, we routinely *underestimate* the presence of unconscious perception (Michel, *Ibid*.). Unconscious perception is powerful, facilitating abilities and actions that one might have otherwise thought would be possible only under conscious perception. For example, in an especially well-known study by Marshall and Halligan (1988), a patient P.S. suffering from unilateral neglect, consistently chose a non-burning house over a burning house as the house she would prefer to live in, despite the engulfed portion of the burning house being placed within her scotoma and only unconsciously perceived. Despite lacking conscious experience of the house, information about the house seems to inform P.S.'s decision-making. Indeed, Berger, Nanay, and Quilty-Dunn (2018) argue that nonconscious perceptual states – like those P.S. has of the engulfed portion of the house – play rational and justificatory roles.

From all of this it seems reasonable to infer that we can have *minded entities* – subjects who perceive and sense and think about the world – whose mental states are not just frequently nonconscious, but that can't be in conscious states at all. These entities might themselves be conscious in some other sense – they might be 'creature conscious' insofar as

[15]For general discussion, see Berger (2014). For evidence from blindsight, unilateral neglect, and masked priming, see Weiskrantz (1997) and Kouider and Dehaene (2007). This is not to say that nonconscious perception is uninformedly accepted. See Phillips (2018) for a different view.

they are awake and responsive to stimuli (Rosenthal 2005) – but they would not have *phenomenally* conscious experiences of the world or themselves.

Such are our prefatory points. Schematically, the mind-first strategy looks like this:

> *The Mind-First Strategy:*
>
> 1. **Step One** (the individuation step): Establish that certain processes are mental in nature and posit an individuation relationship *R* where, for any two mental states M1 and M2, M1 and M2 belong to the same mind iff *R*(M1, M2).[16]
>
> 2. **Step Two:** Ascertain whether *R*(M1, M2), and count the resultant number of minds, treating mind as proxies for subjects.
>
> 3. **Step Three**: (the consciousness step): Employing some criterion for consciousness, ascertain whether the mental states of these subjects are, or can be, conscious.

Notice that a proponent of the mind-first strategy has her work cut out for her in several ways that the proponent of the duplication strategy does not. First, she needs a theory of mindedness: she needs to tell us when something is mental, and when something possesses a mind. Second: she needs to tell us the difference between conscious and nonconscious states. The duplication strategy requires neither of these things. This is because of the duplication principle. It gets us consciousness without appealing to a theory of the difference between nonconscious and conscious mental states (e.g. global workspace theory, higher-order theories), or any view about consciousness's function. And if something has conscious states, it *ipso facto* has mental states. While the mind-first strategy may be applicable to various MV candidates of interest, any iteration of it will require numerous substantive assumptions to even get off the ground.

Before we press on to see how the mind-first strategy might work for the octopus, it will be helpful to have at least some working hypotheses for Step One.

So: what is the lower bound for mentality? In terms of sophistication, we might imagine that *de se* thought occupies one end of the spectrum, and some basic form of non-conceptual perception occupies the other. When it comes to MV, we certainly want something on

[16]Beyond giving different accounts of the *R* relation, different iterations of the mind-first approach might adjust Step One in other ways. For instance, one could try to get by with only a conditional: If M1 and M2 belong to the same mind, then *R*(M1, M2).

the latter end, since there is no demand that MV candidates house multiple cognitively complex subjects. Just so, Carls-Diamante (2019, 465) suggests that the minimal bound for a *cognitive system* is one has causally related 'constituent elements that enable it to carry out sensorimotor coordination operations'. The term 'cognitive system' has a bit of a sub-personal ring to it, which wouldn't help; when we talk about mental states – conscious or not – we are not talking about (say) edge detection in the early visual system.[17] This issue is serious, but not one that can be broached in any detail. Our main point is methodological: it is incumbent on the proponent of the mind-first strategy to take a stand here.[18]

Suppose we simply grant Carls-Diamante that sensory-motor coordination is the lower bound of mentality. There is still the other part of Step One: how do we individuate minds? Since the mind-first approach distinguishes minds from conscious minds, and only counts the latter by first counting the former, it cannot employ relations of p-unity to answer this question. P-unity obtains between conscious states. Thus, counting minds in abstraction from whether minds have conscious states precludes counting via p-unity.

An alternative is to count minds by counting *independent* cognitive systems. Not all cognitive systems are independent – many are mere *subsystems*. The difference resides in their functional autonomy and anatomical boundaries (Carls-Diamante 2019, 466). Independent systems are sufficient for producing a distinctive kind of cognitive output; subsystems may contribute to this effort but are not sufficient on their own. And at least in biological organisms, subsystems are also typically anatomically inseparable from the systems of which they are a part. Take the human visual cortex. It has a well-defined functional role (processing visual information) and is also well-defined anatomically. However, it's lacking in autonomy: if the visual cortex were somehow excised, it would no longer function to process visual information. The visual cortex is a bit like the circulatory system. It's not 'detachable' as a self-contained 'module' that can still serve its function.[19] For this reason, the visual cortex is better viewed as a subsystem of the human neurocognitive system.

[17]Many thanks to Elizabeth Schechter for this point.
[18]To put the point a bit more stridently: what the mind-first strategist is taking a stand on is *the mark of the mental*. This is a massively complicated matter. For just one view on this issue, according to which the mark of the mental is *intentionality*, see Crane (2008).
[19]We owe the 'detachable modules' language to Sidney Carls-Diamante.

Another, arguably better option, is to appeal to the distinction between *direct* and *indirect* mental state interaction. On this option, *R* is a kind of causal relationship. As Schechter (2018, 95) puts it, we can individuate minds with an analysis 'of intramind interactions (interactions between the mental states of a common mind) and of intermind interactions (interactions between the mental states of different minds)'. It's familiar that perceptual states can cause belief states. If you see a brown dog, you will come to believe that there is a brown dog before you. But beliefs can cause perceptual states too. We might think that beliefs can directly cause perceptual states, as in cognitive penetration. Other cases are *indirect,* and rather mundane: if you believe there is beer in the fridge (and want some), this will cause you walk over to the fridge and perceive its contents. Here the belief-perception interface is mediated by various intervening actions and further sensory states. The idea is that intermind interaction is always *exclusively* indirect (Schechter 2018, 95). This is an *architectural* storey about how to individuate minds.[20]

Again, the aim here isn't to settle these questions, but to see how the mind-first strategy might be filled in, and how it might approach MV. It's enough that the answers to the questions aren't unreasonable. With them in place, the mind-first approach can be put to work in perhaps the most interesting of nonhuman test cases for MV: the octopus. We'll begin (§3.1) by rehashing the evidence adduced by Carls-Diamante that octopuses house multiple minds. We'll then (§3.2) take up her claim that octopuses may house multiple *conscious* minds. A quick note: we will frequently put matters in terms of whether the octopus 'arms comprise independent minds' or whether the 'arms have experiences' or whether there are 'arm-based experiences'. Remember: this isn't necessarily to say that the arms *are* subjects of experience; it's to say that we can track, and so count, subjects by counting the arms. If one likes, one can substitute 'peripheral nerve bundles' or 'axial nerve cords' for 'arms'.

[20]When it comes to split brains, Schechter claims that direct mental state interaction is largely confined to *intrahemispheric* interaction, while interhemispheric mental state interaction is mostly indirect. One point of difference here is that Schechter pairs this observation with the duplication principle; so, the direct and indirect casual relations are between mental states that are presumed to be conscious. That's not so here. But there is no reason why such causal interaction must be between conscious mental states. As discussed earlier, if it's plausible than nonconsciously perceiving a flaming house can justify nonconscious beliefs about the house, then it's plausible that these states stand in a causal relationship.

3. A test case for MV: the octopus

3.1. Multiple minds

The octopus arm contains most of the neurons of the octopus's nervous system (350 million of 500 million) of which it is a constituent, subserving specific functional roles. In particular, the octopus arm has sensory-motor responsibilities. Equipped with an immense number of sensory receptors tracking tactile, chemical, and motor information, octopus arms are capable of impressive sensorimotor operations, almost all of which are offloaded from the central brain – a brain that appears to be incapable of supporting *somatotopic representation*, or a point-for-point mapping of the body (Zullo et al. 2009). While the brain issues global motor commands, these must be 'supplemented or filtered' by arm-based sensorimotor information if the octopus is to set spatial parameters (Zullo et al. 2009).

Much of this is due to the octopus' lack of a rigid skeleton: without joints, there are no somatic landmarks from which spatial information about the body can be mapped (Sumbre et al. 2005). Coupled with the flexibility of the octopus's body, this makes motor control challenging. To solve this, Carls-Diamante points out that the octopus employs *stereotypic motor programmes*: motor patterns that store the spatial details of commonly occurring natural movements (Sumbre et al. 2001), 'specify [ing] the set of spatial coordinates the arm must occupy in order to perform a certain action'. Stereotypic motor programmes like *fetching* – a goal-based movement whose success depends on precisely where in the arm the elbow-like medial bend is former – are stored within the arm nervous system and not in the central brain (Carls-Diamante 2019, 473–474).

The peripheral nervous system of the arm is a *cognitive* (and so mental) system. That sensorimotor operations are carried out by the arms is no reason to deny their status as cognitive. Here Carls-Diamante (2019, 476) reminds us of Clark and Chalmers (1998, 8) *parity principle*, which says that if 'a part of the world functions as a process which, were it done in the head, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world is … part of the cognitive process'. If processes outside the body can count as cognitive, processes outside the central brain can too.

Let's just grant, then, the arm-based states are properly mental (surely, if they are cognitive, they are mental). Let M1 be a central brain-based

mental state and let M2 be an arm-based mental state. The next part of Step One from the mind-first strategy asks this: where *R* is an individuation relation, is it the case that *R*(M1, M2)? While this is a bit challenging to pin down in a precise way, we saw that for Carls-Diamante, *R* is something like this: *is realised by the same independent system*. And since, according to Carls-Diamante, it's *not* the case that *R*(M1, M2), it follows that M1 and M2 belong to different minds, i.e. different subjects.[21] Some circularity looms here, though hopefully it is sufficiently moderated by remembering that 'independent system' is analysed in terms of functional autonomy.

Why does Carls-Diamante think that ¬[*R*(M1, M2)]? Two reasons. First, beyond the initiating coarse-grained motor command from the central brain, the octopus arm is fully responsible for its fine-grained sensorimotor coordination operations. The octopus arm is, to this extent, autonomous. Second, the octopus arms retain their sensorimotor capacities even when amputated or the arm's connections to the central brain are severed.

The amputation-based results are admittedly remarkable (Gutfreund et al. 2006; Rowell 1963; Sumbre et al. 2001). Yet notice that Carls-Diamante's individuation criterion essentially seems to imply this: an *intact* octopus houses nine minds (and so nine subjects) because *if* its arms were severed, they would continue to carry out their sensory-motor functions. This seems analogous to: an *intact* human houses two minds (and so two subjects) because *if* its corpus collosum were severed, there would be two p-unified EPs. Granted, some have claimed that non-split-brain humans house two conscious subjects just like split-brain humans. But there is certainly something curious about inferring what the default case is like (octopuses 'in the wild') from how their parts operate in such non-default cases. What we really want – and what seems to be lacking from Carls-Diamante's individuation criterion – is reason to think that the arms comprise independent minds *when the octopus is intact*.[22]

[21] There is a suppressed premise here, which is an analogue of **Ownership**. Call it **Ownership***: all mental states are (non-derivatively) had by a subject. **Ownership*** is technically needed as establishing that M1 and M2 belong to *different* minds requires establishing that M1 and M2 *belong to (are had by) minds*. We have only stipulated that (e.g.) M2 is an 'arm-based' mental state.

[22] Carls-Diamante (2019, 478, emphasis in original) hints at this concern, but then seems to concede the point, saying that '[i]t is reasonable to suppose that there are multiple systems present within the octopus but that they are subsumed under the central, brain-based system *when the animal is intact*'. One might wonder here whether subsumption under the central brain just means there is not enough autonomy (when intact) to justify counting the arm-based system as a distinct mind(s). We take it that this is part of what Schechter (2022) is getting at: counting minds is not just about

Another option is to appeal to Schechter's individuation criterion and ask whether the behavioural independence evidences a lack of *direct causal interdependence* between the states of the peripheral nervous system and the states of the central brain. On the one hand, this criterion has the benefit of not relying on amputation. But on the other, when the octopus *is* intact, there *does* seem to be brain-to-arm direct interaction. Gutnick et al. (2011), for example, have shown that an octopus can use visual information (available just to the central brain) to guide a single arm's movements through a maze. Although we can't, at this stage, assume that the states of the arm are conscious – and so can't assume that this is evidence of a-unity between the *experiences* of the central brain and those of the arms – it does suggest that peripheral arm location information and visual information (whether conscious or not) can be a-unified and thus be employed in controlling goal-directed movements.

Of course, while MV was framed in terms of default conditions – such that, at any given time, one can point in the direction of an (intact, living) octopus and point (in the present parlance) to nine independent cognitive systems and, therefore, nine minded subjects – one might instead think, that, in amputation cases, the octopus fractures into nine minded subjects. This would look a bit like *fission*, as when a pre-fission amoeba ceases to exist upon splitting, with two new amoebas come into existence. Maybe one minded subject (the intact octopus) ceases to exist and, upon amputation, the nine new minded subjects come into existence. Which model best fits the facts is hard to say, but if fission is the right way to view things, it would be a claim different from MV.[23]

3.2. Multiple conscious minds

For sake of argument though, let's grant the arms comprise independent cognitive systems – and so independent minds, and so independent subjects – under default conditions. We can now ask whether the arm-based independent cognitive systems (minds, subjects) are (independently) conscious. On a mind-first approach, this is a separate and complicated step. How might it be carried out?

finding systems that are sufficiently complex in each other's absence (as during amputation), but also sufficiently autonomous *in each other's presence*.

[23]Godfrey-Smith (2020, 179–181) also suggests a functional autonomy-based approach with his *fast switching model*. Somewhat reminiscent of 'contextualist' models (e.g. Tye 2003), the rough idea is that there is a single subject when the brain exerts central control, but two otherwise, with the arm-network comprising one together. Ultimately though, Godfrey-Smith (2019, 4) says that the octopus is unlikely to contain multitudes, be it nine or two.

Consider what we'll call *the structural argument*. It's tempting to think that structural facts about an animal's nervous system at least partially fixes the structure of that animal's consciousness.[24] Thus, given that the octopus has a distributed nervous system, this structural fact should show up experientially. Carls-Diamante (2017, 1278) expands on this idea:

> The great degree of functional self-sufficiency of the octopus arm, and the sparse connections between the central and peripheral components of the nervous system indicate that sensorimotor information from the arms undergoes extensive local processing before being sent to the brain. Presupposing neuralization [i.e. that consciousness is neurally grounded], this motivates the notion that octopus arms may support local phenomenally conscious fields that are independent of the main field whose substrates are found in the brain. That is to say, there is reason to believe that the octopus arm contains neural mechanisms that are necessary—and perhaps sufficient—for generating and supporting a *persisting capacity* for subjective experience. If so, then the sensorimotor responses of an amputated octopus arm to stimulation would be accompanied by corresponding subjective experience, so that there would be something it is like to be an octopus arm.

The structural argument can be summarised like this:

O1. The octopus's nervous system is highly distributed and decentralised across its central brain-based and peripheral arm-based components.
O2. The octopus's arms process sensorimotor information in a way that is sufficiently autonomous from the central brain.
O3. If O1 and O2, then the neural machinery in the octopus's arms support conscious states that are independent from the ones supported by the central brain.
O4. Thus, the neural machinery in the octopus's arms support conscious states that are independent from the ones supported by the central brain.

As it stands, it's unclear how O4 gets us MV. There is a sense in which V4, in the human fusiform gyrus, 'support[s] conscious states that are independent from the ones supported by the' olfactory bulb. The former is critical for the formation of visual representations whereas the olfactory bulb is critical for the formation of olfactory representations. Yet we wouldn't conclude from this that V4 and the olfactory bulb track distinct subjects of experience. What matters is not that different pieces of neural machinery 'support' different kinds of conscious states.

[24]Carls-Diamante (2017, 1271) calls this the *isomorphism thesis*.

What matters is that the arms and the central brain each count as independent cognitive systems in their own right, each of which serves as a *sufficient* local neural basis for conscious experience. Read that way, and allowing that independent cognitive systems can serve as proxies for subjects, O4 would indeed get us MV.

What sort of MV? One option is a two-subjects model. On this view, there are only two subjects, one for the central brain, and one for the arms as a whole. The motivation for this is the interbrachial commissure, which is a neural structure that might serve as an integration mechanism across the eight arms (Carls-Diamante 2017, 1284).

However, the interbrachial commissure doesn't immediately get us two subjects. For consider: while Carls-Diamante (2017, 1270) seems to endorse the left-to-right direction of S-unity ⇔ P-unity, i.e., S-unity ⇒ P-unity, it's open to her to reject S-unity ⇒ P-unity's converse:

> **P-unity ⇒ S-unity**
>
> If E1 and E2 are p-unified at time *t*, then E1 and E2 are (non-derivatively) had by the same subject S at *t*.

Denying **P-unity ⇒ S-unity** permits *between-subject-unity*.[25] This is not easy to envision. The point though is that while the interbrachial commissure would still play the integrative role of p-unifying sensorimotor experiences across the arms, each arm would nonetheless track a numerically distinct subject.[26] The result would be a nine-subjects model.

In fact, it's also open to Carls-Diamante to reject **S-unity ⇒ P-unity** (and by extension, the stronger **S-unity ⇔ P-unity)**. This is critical: recall that the mind-first strategy settles the individuation question – how many subjects are there? – prior to setting whether those subjects are conscious. There is no need to appeal to **S-unity ⇔ P-unity** or **S-unity ⇒ P-unity** to individuate subjects, as on the duplication strategy. Thus, for all that's been said, there could be no *within-subject-unity*. Each arm-based subject could have p-disunified experiences.

So much for interpretive matters. How does the argument fare? The immediate problem is that O3 is false. Recall that what made the duplication strategy appropriate for the split-brain case was that we already had every reason to believe that a single human hemispheric system was on its own sufficient for an experiencing subject. We cannot say anything

[25]For a defence of between-subject-unity, see Roelofs (2016).

[26]We say 'p-unifying *sensorimotor experiences*' because, if the arms have experiences at all, they would be different from central brain-based ones and limited in kind, perhaps to chemotactile and proprioceptive information. See Carls-Diamante (2017, 1280–1281).

analogous about octopuses' arms, however. If we somehow *already* had evidence that the sensorimotor processing found in the arm-based cognitive systems were conscious, then maybe O3 would be justified. But whether we do – more on this momentarily – is a separate issue. That the arms perform sensorimotor processing won't itself settle things, since sensorimotor processing, like anything else, can be and often is nonconscious.

The appeal to functional autonomy thus seems to be doing double duty: first, to help establish that the arms constitute independent cognitive systems; second, to help establish that there is arm-based consciousness that's distinct from central brain-based consciousness. The mind-first strategy lays bare the lacuna in this reasoning. We can't move from the former step to the latter without the intermediary one of ascertaining whether the arm's mental states (assuming they are mental) are conscious states. Indeed, Carls-Diamante seems to acknowledge as much elsewhere when she says that '*if* the brain and the arms can generate local conscious fields, the issue arises as to whether subjective experience in an octopus would be integrated or unified, given the sparseness of interactions between the components of its nervous system' (Ibid. 1273, emphasis added). The question is whether the antecedent of this conditional holds.[27]

So, the structural argument fails. But with the mind-first strategy in view, the fix is in the offing. On the assumption that the arms comprise independent cognitive systems, whether the arms are conscious will depend on the psycho-functional, informational, and neurological features *of those constitutive independent systems* precisely because those systems are independent. In this way, we can ask whether they are conscious in much the same way as we would ask this question of anything else, be it eagles or earthworms.

Consider two approaches to studying nonhuman consciousness from Birch (2022): the *theory-heavy* and *theory-light* approaches.[28] The mind-first strategy can be paired with either. On a theory-heavy approach, we start with a theory of consciousness validated in humans and ask what that theory says, or could say, about nonhumans. For instance, on global-neuronal workspace ('GNW') theory (e.g. Dehaene 2014), mental states are conscious if and only if they are 'in' the GNW and thus available to a

[27] van Woerkum's (2020) takes issue with Carls-Diamante's assumption that the structure of octopus consciousness should be investigated from a neurophysiological and 'internalist' perspective (as evidenced by her neuralization and isomorphism claims). For van Woerkum, the structure of experience is to be explained not in terms of neurophysiology, as Carls-Diamante seems to think, but in terms of 'active' sensorimotor integration. Our criticism differs since, in effect, we are *granting* Carls-Diamante's broadly internalist assumptions, and insisting that the argument still fails.

[28] We ignore the *theory-neutral approach* (Tye 2016) as we agree with Birch that it is a non-starter.

range of consumer systems, including those involved in reporting, planning, reasoning, and decision-making. We would then ask: do octopuses have a GNW? And it turns out that they very well might. Mather (2008) speculates that octopuses must have a 'central processor' akin to that posited by GNW theory due to octopuses' ability to learn and coordinate various visual and tactile cues. So, if the possession of a GNW is sufficient for consciousness, then octopuses have conscious states. Yet this doesn't show that the arm-based mental states (assuming they have them) are conscious: for that to be true, each arm-based system would have to non-derivatively have *its own* GNW.[29] After all, if (somehow) *my* mental states were in *your* GNW, that wouldn't make *me* have conscious experiences.

Or consider higher-order ('HO') theories. It might seem like HO theories are a poor choice for the theory-heavy approach insofar as they are cognitively demanding theories, even more demanding than the GNW theory. Higher-Order Thought ('HOT') theory – arguably the most prominent HO theory – might seem to rule out much or even all nonhuman consciousness from the outset. HOTs are thoughts, which are conceptual states, and the conceptual sophistication of nonhuman animals is arguably limited.

Now technically, not all HO theories appeal to HO states that are conceptual in this way (e.g. Gottlieb 2022). And even HOTs are typically theorised to be rudimentary thoughts. So HOT theory leaves open the possibility that many nonhuman animals can have HOTs (e.g. Berger and Mylopoulos 2024; Rosenthal 2005, 40). Moreover, it's been argued that octopuses *do* have concepts anyway (Mather 2008). So, the only question then would be what those concepts are. But as before, the same issue arises: that octopuses can have consciousness-constituting higher-order states leaves open whether the arm-based subjects nonderivatively do. What we want is not just evidence of independent arm-based sensory-motor processing, but evidence that *each arm-based subject* was somehow aware of its mental states that constituted this sensory-motor processing. And again, we have no evidence for this.

Finally, consider integrated information theories (e.g. Tononi et al. 2016), or 'IIT'. It is not straightforward to give a simple statement of IIT. This is for several reasons (Pautz 2019) that we need not belabor. For

[29]Is this an absurd notion? Not necessarily. Grasso (2014), for example, argues that the octopus has two brains, one of which is the *brachial plexus*, which is the networked ensemble of axial nerve cords. If there is evidence that this network is capable of not just information processing (which seems right), but *the right sort of information integration* prior to any informational connection it may have to another GNW in central brain, then it may well be that the arm-based system has its own GNW. (Imagine two connected, but distinct, routers.) This might only get us two subjects within an octopus, but that is of no matter insofar as MV goes.

present purposes, Michel and Lau's (2020) distinction between *empirical IIT* and *fundamental IIT* will suffice. According to empirical IIT, the right sort of integrated information is a *marker* (i.e. reliable indicator) of consciousness. According to fundamental IIT, consciousness is *identified* with a specific form of integrated information. One could endorse empirical IIT without endorsing fundamental IIT, and as Michel and Lau (2020, 7) note, evidence in favour of empirical IIT could just as well be evidence for other theories of consciousness, like GNW theory. Indeed, Empirical IIT is not really a *theory* of consciousness insofar as it purports to tell us the difference between conscious and nonconscious mental states in the way GNW and HO theories do. Thus, its applicability for the theory-heavy methodology is questionable. So, empirical IIT will be set aside.

Fundamental IIT does not have these problems. It does, alas, have many other problems, not least of which is the dubiousness of its purported axiomatic foundations (Bayne 2018). One might think that if there is any theory on which the arm-based subjects are conscious, it's fundamental IIT. After all, its proponents are frequently upfront about how the theory ascribes consciousness to *nearly everything* (Tononi and Koch 2015, 11), including systems that are much simpler than the octopuses' peripheral nerve bundles. That said, while sensorimotor information is certainly processed and integrated in the arms prior to its transmission to the central brain, IIT identifies consciousness with only the *local maximum* of integrated information ('maximal Φ'). Thus, for all we know, the sensorimotor processing in the arms (when intact) lies outside the region of maximal Φ, in which case it would be nonconscious. Ultimately then, the position we currently face is the same as with GNW and HO theories: we have no evidence that, on IIT, the arm-based subjects have conscious states.

So much for the theory-heavy approach. As an alternative, the theory-light approach attempts to ascertain whether the arm-based systems are conscious without full-blown theoretical commitments:

> A theory-light approach avoids committing to a fully specified theory of consciousness in humans, but it does not eschew theoretical commitments entirely. Instead, it commits to a broad hypothesis about the relation between phenomenal consciousness and cognition that is compatible with a wide range of more specific theories. (Birch 2022, 140)

The 'broad hypothesis' says that '[p]henomenally conscious perception of a stimulus facilitates, relative to unconscious perception, a cluster of cognitive abilities in relation to that stimulus' (Birch 2022). For any way of detailing what cognitive abilities might be in this cluster, we would

then to ask whether the octopuses' arms have those abilities. These abilities would then serve as markers of consciousness. One such marker discussed by Birch is *rapid-reversal learning*, wherein a creature can learn a relationship between two stimuli and then learn the opposite relation. Octopuses seemingly are capable of rapid-reversal learning (Bublitz, Dehnhardt, and Hanke 2021), but we don't have evidence that the arm-based systems can do this individually.

It might be objected that we don't need theories or markers. Rowell (1963) found that, post amputation, pricking an octopus arm with a pin resulted in flinching of the skin and the arm moving away from the direction of the stimulus. This, one might suggest, is direct evidence that the arm-based cognitive systems can harbour pain states. Thus, they are conscious.

The conclusion does not follow. Pain, like any other mental state, can be nonconscious and still drive behavioural responses (Aguilera 2020; Pereplyotchik 2017; Reuter and Sytsma 2020). So, the third step in the mind-first approach – moving from minds to conscious minds – cannot be skirted by adverting to some special subclass of mental states that are invariably conscious. There aren't any. Further, Crook (2021) argues that noxious sensory information is not processed in the arms, but in the central brain. When acetic acid is injected into an arm, pain may be felt *in* the arm but not *by* the arms. But the latter is what we need for MV. Perhaps then what's happening in amputation cases is not pain, let alone conscious pain. After all, paraplegic humans still exhibit the withdrawal reflexes in their lower limbs despite not being able to feel anything below their spinal lesions (e.g. Key 2016, 4). As such, that the arms display withdrawal behaviour need not indicate that they are distinct subjects undergoing experiences of conscious pain.

4. Conclusion

The primary focus of this paper has been methodological. We have not argued that MV is false. What we have argued is that investigating MV, and the counting question it begets, requires sensitivity to two very different strategies: the duplication and mind-first strategies. The duplication strategy – though avoiding certain contentious questions about the nature of mentality and consciousness – is straightforwardly inapplicable in the case of the octopus. By contrast, while the mind-first approach *is* applicable, it is perilous in ways the duplication strategy is not. (Though that's not to say that the duplication strategy can be run carefree.) Whereas it's reasonable to think human split-brain patients have multiple

loci of consciousness associated with each hemisphere – even if these loci ultimately don't track distinct conscious *subjects* – we can't say even that much for the octopus. If the claim that the octopus contains multitudes is eventually substantiated, it will be after much more work, and the duplication strategy won't deliver it.[30]

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Joseph Gottlieb http://orcid.org/0000-0001-9014-8487
Bob Fischer http://orcid.org/0000-0001-9605-393X

References

Aguilera, B. 2020. "Nonconscious Pain, Suffering, and Moral Status." *Neuroethics* 13 (3): 337–345. https://doi.org/10.1007/s12152-020-09430-y

Baker, L. R. 2000. *Persons and Bodies: A Constitution View*. New York: Cambridge University Press.

Bayne, T. 2008. "The Unity of Consciousness and the Split-Brain Syndrome." *Journal of Philosophy* 105 (6): 277–300. https://doi.org/10.5840/jphil2008105638.

Bayne, T. 2010. *The Unity of Consciousness*. Oxford: Oxford University Press.

Bayne, T. 2018. "On the Axiomatic Foundations of the Integrated Information Theory of Consciousness." *Neuroscience of Consciousness*. https://doi.org/10.1093/nc/niy007.

Bayne, T., and D. J. Chalmers. 2003. "What is the Unity of Consciousness?" In *The Unity of Consciousness: Binding, Integration, and Dissociation*, edited by A. Cleeremans, 23–58. Oxford: Oxford University Press.

Berger, J. 2014. "Mental States, Conscious and Nonconscious." *Philosophy Compass* 9 (6): 392–401. https://doi.org/10.1111/phc3.12140.

Berger, J., and M. Mylopoulos. 2024. "HOTT and Heavy: Higher-Order Thought Theory and the Theory-Heavy Approach to Animal Consciousness." *Synthese* 203 (98). https://doi.org/10.1007/s11229-024-04529-8.

Berger, J., B. Nanay, and J. Quilty-Dunn. 2018. "Unconscious Perceptual Justification." *Inquiry* 61 (5-6): 569–589. https://doi.org/10.1080/0020174X.2018.1432413.

Birch, J. 2022. "The Search for Invertebrate Consciousness." *Nous* 56 (1): 133–153. https://doi.org/10.1111/nous.12351.

Birch, J., A. K. Schnell, and N. S. Clayton. 2020. "Dimensions of Animal Consciousness." *Trends in Cognitive Sciences* 24 (10): 789–801. https://doi.org/10.1016/j.tics.2020.07.007.

[30]For comments and discussion, we thank Sidney Carls-Diamante, Elizabeth Schechter, Alex Schnell, and several anonymous referees. We also thank Rethink Priorities for supporting earlier versions of this research.

Block, N. 1995. "On a Confusion about a Function of Consciousness." *Brain and Behavioral Sciences* 18 (2): 227–247.

Block, N. 2007. "Consciousness, Accessibility, and the Mesh Between Psychology and Neuroscience." *Behavioral and Brain Sciences* 30 (5-6): 481–499. https://doi.org/10.1017/S0140525X07002786.

Bradford, G. 2023. "Consciousness and Welfare Subjectivity." *Nous (Detroit, Mich)* 57 (4): 905–921. https://doi.org/10.1111/nous.12434.

Bublitz, A., G. Dehnhardt, and F. D. Hanke. 2021. "Reversal of a Spatial Discrimination Task in the Common Octopus (Octopus Vulgaris)." *Front. Behav. Neurosci* 15: 614523. https://doi.org/10.3389/fnbeh.2021.614523.

Carls-Diamante, S. 2017. "The Octopus and the Unity of Consciousness." *Biology and Philosophy* 32 (6): 1269–1287. https://doi.org/10.1007/s10539-017-9604-0.

Carls-Diamante, S. 2019. "Out on a Limb? On Multiple Cognitive Systems Within the Octopus Nervous System." *Philosophical Psychology* 32 (4): 463–482. https://doi.org/10.1080/09515089.2019.1585797.

Carls-Diamante, S. 2022. "Where is it Like to be an Octopus?" *Frontiers in Systems Neuroscience* 16: 840022. https://doi.org/10.3389/fnsys.2022.840022.

Carruthers, P. 2005. "Why the Question of Animal Consciousness Might not Matter Very Much." *Philosophical Psychology* 18 (1): 83–102. https://doi.org/10.1080/09515080500085411

Clark, A., and D. Chalmers. 1998. "The Extended Mind." *Analysis* 58 (1): 7–19. https://doi.org/10.1093/analys/58.1.7

Crane, T. 2008. "Reply to Nes." *Analysis* 68 (299): 215–218. https://doi.org/10.1111/j.1467-8284.2008.00740.x

Crook, R. J. 2021. "Behavioral and Neurophysiological Evidence Suggests Affective Pain Experience in Octopus." *iScience* 24 (3): 102229. https://doi.org/10.1016/j.isci.2021.102229.

Crummett, D. 2022. "What if We Contain Multiple Morally Relevant Subjects?" *Utilitas* 34 (3): 317–334. https://doi.org/10.1017/S0953820822000115.

Dainton, B. 2008. *The Phenomenal Self*. Oxford: Oxford University Press.

Dehaene, S. 2014. *Consciousness and the Brain: Deciphering How the Brain Encodes Our Thoughts*. New York: Viking Press.

Dretske, F. 1995. *Naturalizing the Mind*. Cambridge: MIT Press.

Godfrey-Smith, P. 2016. *Other Minds: The Octopus, the Sea, and the Deep Origins of Consciousness*. New York: Farrar, Straus and Giroux.

Godfrey-Smith, P. 2019. "Octopus Experience." *Animal Sentience* 26 (18). https://doi.org/10.51291/2377-7478.1503.

Godfrey-Smith, P. 2020. *Metazoa: Animal Life and the Birth of the Mind*. New York: Farrar, Straus and Giroux.

Godfrey-Smith, P. 2021. "Integration, Lateralization, and Animal Experience." *Mind and Language* 36 (2): 285–296. https://doi.org/10.1111/mila.12323

Gottlieb, J. 2022. "The Higher-Order Map Theory of Consciousness." *Australasian Journal of Philosophy* 176 (1): 1–20.

Gottlieb, J., and B. Fischer. 2024. "The Ethical Implications of Panpsychism." *Australasian Journal of Philosophy*, 1–15. https://doi.org/10.1080/00048402.2024.2350708.

Grasso, F. W. 2014. "The Octopus with Two Brains: How Are Distributed and Central Representations Integrated in the Octopus Central Nervous System?" In *Cephalopod Cognition*, edited by Anne-Sophie Darmaillacq, Ludovic Dickel, and Jennifer Mather, 94–122. Cambridge, UK: Cambridge University Press.

Gutfreund, Y., H. Matzner, T. Flash, and B. Hochner. 2006. "Patterns of Motor Activity in the Isolated Nerve Cord of the Octopus Arm." *The Biological Bulletin* 211 (3): 212–222. https://doi.org/10.2307/4134544.

Gutnick, T., R. A. Byrne, B. Hochner, and M. Kuba. 2011. "'Octopus Vulgaris Uses Visual Information to Determine the Location of its arm.'." *Current Biology* 21 (6): 460–462. https://doi.org/10.1016/j.cub.2011.01.052.

Key, B. 2016. "Why Fish Do Not Feel Pain." *Animal Sentience* 1 (3). https://doi.org/10.51291/2377-7478.1011.

Kouider, S., and S. Dehaene. 2007. "Levels of Processing During Non-Conscious Perception: A Critical Review of Visual Masking." *Philosophical Transactions of the Royal Society B* 362 (1481): 857–875. https://doi.org/10.1098/rstb.2007.2093.

Kriegel, U. 2019. "The Value of Consciousness." *Analysis* 79 (3): 503–520. https://doi.org/10.1093/analys/anz045

Levy, N. 2024. "Consciousness Ain't All That." *Neuroethics* 17 (2): 1–14. https://doi.org/10.1007/s12152-024-09559-0

Marshall, J. C., and P. W. Halligan. 1988. "Blindsight and Insight in Visuo-Spatial Neglect." *Nature* 336 (6201): 766–767. https://doi.org/10.1038/336766a0.

Mather, J. 2008. "Cephalopod Consciousness: Behavioural Evidence." *Consciousness and Cognition* 17 (1): 37–48. https://doi.org/10.1016/j.concog.2006.11.006

Mather, J. 2021. "The Case for Octopus Consciousness: Unity." *NeuroSci* 2 (4): 405–415. https://doi.org/10.3390/neurosci2040030.

Michel, M. 2023. "How (Not) to Underestimate Unconscious Perception." *Mind & Language* 38 (2): 413–430. https://doi.org/10.1111/mila.12406.

Michel, M., and H. Lau. 2020. "On the Dangers of Conflating Strong and Weak Versions of a Theory of Consciousness." *Philosophy and the Mind Sciences* 1 (II): 8. https://doi.org/10.33735/phimisci.2020.II.54.

Mill, J. S. 1861/1969. "Utilitarianism." In *The Collected Works of John Stuart Mill*, Vol. X, Essays on Ethics, Religion and Society, edited by J. M. Robson, 203–259. Toronto and London: University of Toronto Press and Routledge & Kegan Paul.

Nagel, T. 1971. "Brain Bisection and the Unity of Consciousness." *Synthese* 22 (3-4): 396–413. https://doi.org/10.1007/BF00413435

O'Conaill, D. 2024. "The Having Objection to Bundle Theories of Subjects of Experience." *The Philosophical Quarterly*, pqae040. https://doi.org/10.1093/pq/pqae040.

Pautz, A. 2019. "What is the Integrated Information Theory of Consciousness?" *Journal of Consciousness Studies* 26 (1-2): 188–215.

Pereplyotchik, D. 2017. "Pain and Consciousness." In *The Routledge Handbook of the Philosophy of Pain*, edited by J. Corns, 210–220. New York: Routledge.

Phillips, I. 2018. "Unconscious Perception Reconsidered." *Analytic Philosophy* 59 (4): 471–514.

Reuter, K., and J. Sytsma. 2020. "Unfelt Pain." *Synthese* 197 (4): 1777–1801. https://doi.org/10.1007/s11229-018-1770-3

Roelofs, L. 2016. "The Unity of Consciousness, Within Subjects and Between Subjects." *Philosophical Studies* 173 (12): 3199–3221. https://doi.org/10.1007/s11098-016-0658-7

Roelofs, L. 2019. *Combining Minds: How to Think about Composite Subjectivity*. Oxford: Oxford University Press.

Rosenthal, D. M. 2005. *Consciousness and Mind*. Oxford: Oxford University Press.

Roth, G., and U. Dicke. 2017. "Evolution of Cognitive Brains: Mammals." In *The Evolution of the Brain, Cognition, and Emotion in Vertebrates*, edited by S. Watanabe, M. A. Hofman, and T. Shimizu, 125–146. Tokyo: Springer. https://doi.org/10.1007/978-4-431-56559-8.

Rowell, C. H. F. 1963. "Excitatory and Inhibitory Pathways in the arm of Octopus." *Journal of Experimental Biology* 40 (2): 257–270. https://doi.org/10.1242/jeb.40.2.257

Schechter, E. 2018. *Self- Consciousness and "Split" Brains: The Minds' I*. Oxford: Oxford University Press.

Schechter, E. 2022. "Reply to Commentators." *Journal of Consciousness Studies* 29 (1-2): 179–203. https://doi.org/10.53765/20512201.29.1.179

Sumbre, G., G. Fiorito, T. Flash, and B. Hochner. 2005. "Motor Control of Flexible Octopus Arms." *Nature* 433 (7026): 595–596. https://doi.org/10.1038/433595a

Sumbre, G., Y. Gutfreund, G. Fiorito, T. Flash, and B. Hochner. 2001. "Control of Octopus arm Extension by a Peripheral Motor Program." *Science* 293 (5536): 1845–1848. https://doi.org/10.1126/science.1060976

Tononi, G., M. Boly, M. Massimini, and C. Koch. 2016. "Integrated Information Theory: From Consciousness to its Physical Substrate." *Nature Reviews Neuroscience* 17 (7): 450–461. https://doi.org/10.1038/nrn.2016.44

Tononi, G., and C. Koch. 2015. "Consciousness: Here, There and Everywhere?" *Philosophical Transactions of the Royal Society B: Biological Sciences* 370 (1668): 20140167. https://doi.org/10.1098/rstb.2014.0167.

Tye, M. 2003. *Consciousness and Persons: Unity and Identity*. Cambridge: MIT Press.

Tye, M. 2009. *Consciousness Revisited. Materialism Without Phenomenal Concepts*. Cambridge: MIT Press.

Tye, M. 2016. *Tense Bees and Shell-Shocked Crabs: Are Animals Conscious?* New York: Oxford University Press.

Vallortigara, G. 2000. "Comparative Neuropsychology of the Dual Brain: A Stroll Through Animals' Left and Right Perceptual Worlds." *Brain and Language* 73 (2): 189–219. https://doi.org/10.1006/brln.2000.2303.

van Woerkum, B. 2020. "Distributed Nervous System, Disunified Consciousness?: A Sensorimotor Integrationist Account of Octopus Consciousness." *Journal of Consciousness Studies* 27 (1-2): 149–172.

Veit, W. 2023. *A Philosophy for the Science of Animal Consciousness*. London: Routledge.

Weiskrantz, L. 1997. *Consciousness Lost and Found: A Neuropsychological Exploration*. Oxford: Oxford University Press.

Zullo, L., G. Sumbre, C. Agnisola, T. Flash, and B. Hochner. 2009. "Nonsomatotopic Organization of the Higher Motor Centers in Octopus." *Current Biology* 19 (19): 1632–1636. https://doi.org/10.1016/j.cub.2009.07.067