Contents

Research Articles

Organon F 31 (3) 2024: 204–216
https://doi.org/10.31577/orgf.2024.31301

ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

Intentions, Commitments, and the Derivation of Implicatures

Matej Drobňák*

Received: 8 April 2024 / Accepted: 30 June 2024

Abstract: In this paper, I focus on a common equivocality in how the content of conversational, especially scalar, implicatures is specified and I argue that there is a substantial difference between the belief specification $\text{BEL}_S(\neg\psi)$ ("The speaker believes that $\neg\psi$") and the content specification $\neg\psi$. The main argument for taking the distinction between the specifications seriously is that, in most cases, both $\text{BEL}_S(\neg\psi)$ and $\neg\psi$ can be derived as the implicatures of the same sentence but they have different consequences for how the hearer plans her future actions and manages expectations about the future actions of the speaker. As I argue further, the commitment-based approaches can provide an explanation of how the content specification is derived in contexts in which the speaker does not have beliefs required for the derivation of the belief specification and because of that they have an advantage over the standard Gricean approach.

Keywords: Commitments; Gricean pragmatics; implicatures; intentions; social-normative pragmatics.

* University of Hradec Králové
https://orcid.org/0000-0001-7113-2543
Department of Philosophy and Social Sciences, nám. Svobody 331/2, 500 02 Hradec Králové 2, Czechia.
matej.drobnak@gmail.com

1. Introduction

If we take a closer look at the literature on scalar implicatures, we can notice a very common equivocality in the way the content of implicatures is specified.[1] On the one hand, implicatures are often specified as the *contents* of beliefs of a speaker. On the other hand, implicatures are also specified as the *beliefs* (or other intentional states) of a speaker itself. To mark the difference better, we can represent the content specification by "$\neg\psi$" ("ψ is false" where ψ is some proposition) and the belief specification by "$\mathrm{BEL}_S(\neg\psi)$" ("The speaker believes that ψ is false"). An example of this equivocality can be found in Carston (1998). When she gives examples of scalar implicatures, she uses both specifications:

(1) a. Bill has got some of Chomsky's papers.
b. The speaker believes that Bill hasn't got all of Chomsky's papers.

(2) a. There will be five of us for dinner tonight.
b. There won't be more than five of us for dinner tonight.

(3) a. X: I like Mary. She's intelligent and good-hearted.
Y: She's intelligent.
b. Y doesn't think Mary is good-hearted.

(4) a. She won't necessarily get the job.
b. She will possibly get the job. (Carston 1998, 179)

In (1b) and (3b), the implicatures are specified as beliefs (or other intentional states) of the speakers, while in (2b) and (4b), the implicatures are specified as the contents of beliefs.

Another example can be found in Sauerland (2004). Sauerland (2004, 369) paraphrases scalar implicatures as having the form "The speaker is certain that ψ is false" or "$\mathrm{K}\neg\psi$" (where the certainty operator K approximates knowledge, and so an intentional state, of the speaker). But when he presents examples of scalar implicatures, he omits the part specifying the intentional states and uses "ψ is false" or "$\neg\psi$" only. Similar examples are

[1] In what follows, I focus mostly on the discussions about scalar implicatures as the equivocality is most visible there. But the arguments presented in the paper have broader implications for how conversational implicatures in general can be derived.

ubiquitous in literature and authors often switch from one specification to another arbitrarily.[2]

The general aim of the paper is to argue that the distinction between $\text{BEL}_S(\neg\psi)$ and $\neg\psi$ is substantial and that the two specifications cannot be replaced by each other arbitrarily. The particular aims of the paper are to show that a) the same sentence can carry both $\text{BEL}_S(\neg\psi)$ and $\neg\psi$ as its implicatures, b) $\text{BEL}_S(\neg\psi)$ and $\neg\psi$ cannot be used interchangeably because they have different impacts on the overall conversational situation, and c) the potential impact on the overall situation influences whether $\text{BEL}_S(\neg\psi)$ or $\neg\psi$ is a preferred option in a particular context.

In Section 2, I propose a possible explanation of why the equivocality has been mostly overlooked in literature and I argue that the belief specification should be preferred from the Gricean "intentionalist" perspective. In Section 3, I argue that the belief specification is a preferred option only in specific contexts in which information about the beliefs of the speaker has an impact on the overall conversational situation, especially on the way the hearer plans her own actions and manages her expectations about the actions of the speaker. In Section 4, I argue that the content specification is preferred, and it may even be the only option available, in many other contexts. In Section 5, I argue that commitment-based approaches (Brandom 1994, 2000; Geurts 2019a, 2019b) can provide a way for deriving the content specification in contexts in which the speaker does not have beliefs required for the derivation of the belief specification and so they have an advantage over the standard Gricean approach.

2. Gricean communication and the standard recipe

The equivocality between $\text{BEL}_S(\neg\psi)$ and $\neg\psi$ has a rather simple historical explanation. As a matter of fact, the discussions on implicatures have

[2] A rare example of a paper that makes the distinction explicitly is Franke (2012). Franke distinguishes between base-level implicatures ($\neg\psi$) and strong epistemic implicatures ($\text{BEL}_S(\neg\psi)$) and provides game-theoretic models for deriving both. A general idea behind the models will be briefly discussed later in the paper.

started with Grice (1975, 1978, 1981) and one of the most distinctive components of his views is the intentionalist view of communication. According to this view, communication is a matter of the hearers' recognition and ascription of speakers' intentions (Grice 1957, 1968, 1969).

In general, we can find numerous proposals for how different types of conversational implicatures can be derived in the spirit of the intentionalist view and one of the most specific and the most common proposals is the so-called standard recipe for the derivation of scalar implicatures:

> S has said ϕ.
> *i.* S could have made a stronger claim by saying ψ. Why didn't he do so?
> *ii.* Presumably, it's because S doesn't believe that ψ is true: $\neg \mathrm{BEL_S}(\psi)$.
> *iii.* S has an opinion as to whether ψ is true: $\mathrm{BEL_S}(\psi) \vee \mathrm{BEL_S}(\neg\psi)$.
> *iv.* Between them, (*ii*) and (*iii*) entail $\mathrm{BEL_S}(\neg\psi)$: S believes that ψ is false. (Geurts 2010, 32)[3]

For example, if ϕ = (5a), then we can use the standard recipe to derive (5b).

(5) a. John drank some of the beers.
b. The speaker believes that John did not drink all the beers.

Derivation of scalar implicatures through the standard recipe represents a specific case of the general intentionalist view in which the hearers engage in abductive reasoning about the intentional states of the speakers. As part of the standard recipe, the hearer is supposed to find the best explanation for the speaker's choice of words by making assumptions about the intentional states of the speaker. In particular, she makes assumptions about her beliefs represented by the premises (*ii*) and (*iii*). The belief operator used in the premises naturally passes on to the conclusion, resulting in the belief specification of scalar implicatures.

For someone who accepts the intentionalist view of communication, as the vast majority of researchers in pragmatics do, it might be hard to notice

[3] The argument is implicitly present in Grice (1975, 1978) already. The first explicit version of the standard recipe can be found in Soames (1982). Besides Geurts (2010), we can find the fully explicit form of the argument e.g. Horn (1989), or van Rooij and Schulz (2004).

the equivocality, because she may automatically tend to read the content specification $\neg\psi$ as representing the intentional states of the speaker. And for someone who considers the standard recipe to be a plausible approximation of how scalar implicatures are derived, the belief specification $\text{BEL}_S(\neg\psi)$ should be *the* specification of the content of implicatures, as $\text{BEL}_S(\neg\psi)$ is the actual output of the standard recipe. Taking into account the popularity of the standard recipe, I assume that the belief specification is the generally preferred, though not always explicitly recognized, specification of the content of conversational implicatures.

If this is so, then the equivocality discussed here can be understood as a pardonable simplification. The idea would be that, strictly speaking, $\text{BEL}_S(\neg\psi)$ is the right way to specify the content of scalar implicatures. The content specification $\neg\psi$ is just a shorter, imprecise placeholder for the longer specification used by authors for the sake of brevity.

Another reason for understanding the derivation of scalar implicatures through the lens of standard recipe, and so for seeing $\text{BEL}_S(\neg\psi)$ as the preferred option for the specification of the content of scalar implicatures, is that the standard recipe allows us to make a difference between weak ($\neg\text{BEL}_S(\psi)$) and strong ($\text{BEL}_S(\neg\psi)$) implicatures. This distinction would not be possible to make if the content specification would be accepted as the specification of the content of scalar implicatures.

3. When intentional states matter

In this section, I argue that the distinction between the belief and the content specification is not just a matter of superficial terminological sloppiness. The main reason for taking the distinction seriously is that many sentences can carry both $\text{BEL}_S(\neg\psi)$ and $\neg\psi$ as their implicatures and whether the hearer derives $\text{BEL}_S(\neg\psi)$ or $\neg\psi$ has profound consequences for how she plans and execute her subsequent actions. To be open, I do not see any reason that would prevent the option that the hearers can switch between deriving $\text{BEL}_S(\neg\psi)$ or $\neg\psi$ depending on their momentary interests as most of the conversational situations allow for deriving both $\text{BEL}_S(\neg\psi)$ and $\neg\psi$. However, there are specific situations in which either one or the

other specification is preferred. We can demonstrate this through the example of the sentence (6a).

(6) a. Some civilians left the building before the attack.
b. The speaker believes that not all civilians left the building before the attack.
c. Not all civilians left the building before the attack.

I focus first on a situation in which derivation of $\text{BEL}_S(\neg\psi)$ is preferred. Let us say that intelligence agency discovered a hideout of a terrorist, and the president ordered an attack in order to eliminate the terrorist. Unfortunately, some civilians died during the attack and the president is now facing a trial for homicide. The prosecutor interrogates the president:

Pro.: Did you know anything about the position of civilians before ordering the attack?
Pre.: Some civilians left the building before the attack.

If the hearers are the prosecutor, a judge, and a jury, then the belief specification (6b) is a preferred option for the implicature of (6a), because the intentional states of the president (the speaker) matters for the trial. In particular, information that the president had a specific belief has a profound impact on the decisions and actions of the hearers. If the president believed that all civilians left the building before the attack, then she might be accused of unintentional homicide by the prosecutor and judged for committing a manslaughter by the judge and the jury. If the president believed that not all civilians left the building before the attack, but ordered it anyway, then she could be accused of intentional homicide by the prosecutor and judged for committing a murder by the judge and the jury.

What makes $\text{BEL}_S(\neg\psi)$ the contextually preferred option is that derivation of $\text{BEL}_S(\neg\psi)$ makes it possible for the hearers to make actions that would not be possible if $\neg\psi$ would be derived in this situation. In particular, if $\neg\psi$ would be derived by the hearers, then it would not be possible for them to decide whether the committed crime should be classified as a murder or a manslaughter as (6c) carries no information about the beliefs of the president and so it has no bearing on the intentionality of her actions.

Generally speaking, $\text{BEL}_S(\neg\psi)$ is preferred as an implicature in those situations in which information about the intentional states of the speaker

have consequences for the subsequent actions of the hearers beyond and above the consequences provided by $\neg\psi$. Standardly, these are contexts in which the hearers assess the actions or the character of the speaker in order to plan their own future actions and manage their expectations regarding the actions of the speaker. Deciding whether I should vote for the president in the upcoming elections, whether I want to go with someone to a family vacation, whether I want to start a business with someone, or simply whether I can trust someone are among the paradigmatic examples.

4. When intentional states do not matter

However, situations in which $\text{BEL}_S(\neg\psi)$ is preferred are far from ubiquitous. In other words, contrary to what the standard recipe might suggest, the derivation of $\text{BEL}_S(\neg\psi)$ is in no way universal. As a matter of fact, the intentional states of the speaker might be irrelevant and $\neg\psi$ might be the preferred option for the implicature of (6a) in many other situations. For example, let us say that John's wife is working as a cleaner in the Pentagon. There has been a terrorist attack on the building. When John arrives there, he gets into a conversation with a rescuer:

John: My wife was at work here today.
Res.: What is her job?
John: She is a cleaner.
Res.: Some civilians left the building before the attack.

In this situation, the content specification (6c) is a preferred option for the implicature of (6a), because $\neg\psi$ provides sufficient information for John (the hearer). If all civilians left the building before the attack, then John's wife should be safe and he does not have to worry about her life. This is no longer true if he derives (6c) as the implicature as it leaves the possibility that his wife was in the building during the attack open. Deriving $\neg\psi$ has an impact on how John will plan his subsequent actions, e.g. he might calm down a little, ask for a list of those who did not leave the building before the attack, or where those who are safe are.

Generally speaking, $\neg\psi$ is preferred as an implicature in those situations in which information about the intentional states of the speaker have no

direct consequences for the subsequent actions of the hearer. Standardly, these are situations in which the actions of the hearer are guided only by information provided by $\neg\psi$ and the management of expectations about the future actions of the speaker does not play a role (e.g. one-off encounters). Asking for a tip for a restaurant ("Some restaurants in this area are good") or asking a stranger for directions ("Some buses from this stop go there") are among the paradigmatic examples.

5. How to derive the content specification?

The question I would like to raise now is: If a sentence can carry both the belief specification and the content specification as its implicatures, what is the relation between them with respect to their derivability? From the Gricean perspective, it seems natural to assume that the derivation is serial: the hearer first derives the belief specification and, on the basis of that, she derives the content specification.[4] We can see how this could be the case through the example of a paramedic-patient conversation. Let us say that a paramedic examines a patient who does not feel well and the patient utters (7a).

(7) a. I feel very cold.
b. The patient has hypothermia.
c. The patient believes she has hypothermia.

In this situation, taking into account further symptoms (e.g. shivering), the paramedic (the hearer) preferably derives the content specification (7b) as the implicature, because (7b) is relevant for the planning of her subsequent actions, in particular, for identifying and initiating the proper treatment (e.g. removing wet clothes, providing warm clothes or blankets, providing a sweet beverage). If we assume that the patient has advanced medical knowledge (e.g. she is a doctor), then the paramedic could derive (7b) seri-

[4] A similar suggestion has been proposed by Franke (2012, 7) for scalar implicatures. In his reading, the preferred Gricean option is a serial derivation in which $\neg\text{BEL}_S(\psi)$ is strengthened to $\text{BEL}_S(\neg\psi)$, which may in turn be strengthened to $\neg\psi$.

ally by relying on the maxim of quantity: "Make your contribution as informative as is required (for the current purposes of the exchange)" (Grice 1975, 26).

1. The patient has said that ϕ [= (7a)].
2. There is no reason to suppose that she is not observing the maxims.
3. She could not be doing this unless she thought that ψ [= (8)/(7b)]: $\mathrm{BEL}_S(\psi)$.
4. She knows (and knows that I know that she knows) that I can see that the supposition that she thinks that ψ is required.
5. She has done nothing to stop me thinking that ψ.
6. She intends me to think, or is at least willing to allow me to think, that ψ.
7. Therefore, she has implicated that ψ.

The idea is that (7a) is not as informative as it is required with respect to the goal of the conversation, i.e. for identifying and initiating a treatment, and (8) represents its more informative alternative.

(8) I have hypothermia.

Since the patient used (7a) instead of (8), she could not abide the maxim of quantity unless she *believes* (8). In other words, the serial derivation requires that the paramedic first derives the belief specification (7c) in Step 3 and, only on the basis of that, she derives the content specification (7b) in Step 7.

This serial derivation requires that the patient has an expertise in specific medical knowledge. This assumption, however, certainly does not hold often in the paramedic-patient encounters. In particular, in some other case, the paramedic may reasonably believe (or even know) that the patient does not know what hypothermia is. Despite that, it seems plausible that she derives the same implicature (7b) in such a context as well, because information communicated in (7b) is crucial for planning her subsequent actions (and, arguably, there is no reason to suppose that the paramedic would behave differently in the first and the second scenario).

However, in such a case, the paramedic cannot derive (7b) through the belief specification (7c), because such a derivation would require ascribing the patient the belief that she may not possibly have. This poses a challenge

for the Gricean view of communication. How can the hearer derive a conversational implicature without relying on the considerations about the intentional states of the speaker?

What sounds like an oxymoron from the Gricean perspective may have a perfectly reasonable explanation from another perspective. An especially promising alternative for showing how conversational implicatures could be derived without considerations about the intentional states of the speakers are the commitment-based approaches (Brandom 1994, 2000; Peregrin 2014; Geurts 2019a, 2019b). Setting aside the differences, the commitment-based approaches hold that the primary aim of communication is to establish commitments and, by doing this, to help to coordinate the actions between the speakers and the hearers. In particular, commitments can help the speaker and the hearer to plan their actions in accordance with the expectations induced by the commitments. As Geurts (2019a, 3) puts it, "If Albert promises Brenda to do the dishes, he commits himself to do the dishes, and by the same token Brenda becomes entitled to act on the assumption that Albert will do the dishes".

Interesting feature of the commitment-based approaches is that the establishment of commitments is a social matter and so it does not always rely on the intentions of the speakers. For example, I might become legally married to someone by uttering (9) while being drunk in the wedding chapel in Las Vegas regardless of the fact whether I intend to marry the second person.

(9) I do.

This feature of the view can help us to explain how the paramedic can derive the content specification in the context in which the speaker does not have the beliefs required for the derivation of the belief specification. In particular, the paramedic could use the commitment-based alternative of the serial derivation:

1. The patient has said that ϕ [= (7a)].
2. There is no reason to suppose that she is not observing the maxims.
3. She could not be doing this unless she was committed to ψ [= (8)/(7b)]: $\mathrm{COM_S}(\psi)$.
4. She did nothing to stop me in attributing the commitment to ψ to her.

5. She intends me, or is at least willing to allow me to attribute the commitment to ψ to her.
6. Therefore, she has implicated that ψ.

The idea is that the commitment-based approaches can explain how implicatures are derived in situations in which there are discrepancies between the beliefs of the speakers and the hearers, because whether a commitment is established does not necessarily depend on whether the speaker intends to establish the commitment. In particular, the paramedic can attribute the commitment to (7b) to the patient regardless of the fact whether the patient acknowledges the commitment (Brandom 1994, 194).

What I consider crucial for such attributions is the second part of the assumption in Step 5 of the derivation, i.e. the assumption that the hearer is willing to allow the speaker to attribute the commitment to ψ to her. This assumption can be supported by more general considerations about the division of linguistic labour. In particular, if it is a common ground between the patient and the paramedic that the paramedic has an expertise in medicine, then the patient should acknowledge any commitment attributed to her by the paramedic as the paramedic is an expert on which commitments are actually established in such situations. In other words, the paramedic is entitled to attribute the commitments to the patient and the patient should be willing to undertake the commitments without knowing exactly what she is committing to at the time of uttering the sentence. Since the patient does not know which commitments she undertakes, the paramedic can guide her in how she should behave in accordance with the attributed commitments by giving her instructions and she should be ready to follow these instructions.

From this perspective, we can see the derivation of conversational implicatures as not being based on considerations about the intentional states of the speaker, but as being based on considerations about the commitments established in a particular context. In particular, the reliance on commitments can explain how the content specification could be derived without a previous derivation of the belief specification.

6. Conclusions

In this paper, I discuss the equivocality present in the way the content of conversational, especially scalar, implicatures is specified and I argue that the difference between the belief specification $\mathrm{BEL_S}(\neg\psi)$ and the content specification $\neg\psi$ is not a matter of superficial terminological sloppiness. The main argument for taking the distinction between the belief and the content specification seriously is that both specifications can be derived as the implicatures of the same sentence, and they have different consequences with respect to the hearer and the way she plans her own future actions and manages expectations about the future actions of the speaker. If this is so, then the question arises how the content specification $\neg\psi$ is derived. As I argue, the commitment-based approaches provide an explanation of how the content specification can be derived without making considerations about the intentional states of the speakers by relying on commitments that are attributed to the hearer (but which are not necessarily acknowledged by the hearer). Because of that, the commitment-based approaches have an advantage over the Gricean view in the cases in which the speakers do not have relevant intentional states to communicate an implicature.

Acknowledgements

I would like to thank an anonymous reviewer for the constructive and helpful feedback.

Funding

This work was supported by The Czech Science Foundation (GAČR), grant number 22-05200O.

References

Brandom, Robert. 1994. *Making it Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, Massachisetts: Harvard University Press.

Brandom, Robert. 2000. *Articulating Reasons: An Introduction to Inferentialism*. Cambridge, Massachusetts: Harvard University Press.

Carston, Robyn. 1998. "Informativeness, relevance and scalar implicature." In *Relevance Theory: Applications and Implications*, edited by Seiji Uchida and Robyn Carston, 179-–38. Amsterdam: John Benjamins.

Franke, Michael. 2011. "Quantity Implicatures, Exhaustive Interpretation, and Rationalconversation." *Semantics and Pragmatics*, 4: 1–82. https://doi.org/10.3765/sp.4.1

Geurts, Bart. 2010. *Quantity Implicatures*. Cambridge: Cambridge University Press.

Geurts, Bart. 2019a. "Communication as Commitment Sharing: Speech Acts, Implicatures, Common ground." *Theoretical Linguistics*, 45(1-2): 1–30. https://doi.org/10.1515/tl-2019-0001

Geurts, Bart. 2019b. "What's Wrong with Gricean Pragmatics." In *Proceedings of ExLing 2019*, edited by Antonis Botinis, 1-8. Lisbon: ExLing Society.

Grice, Paul. 1957. "Meaning." *Philosophical Review*, 66(3): 377–88.

Grice, Paul. 1968. "Utterer's Meaning, Sentence-Meaning, and Word-Meaning." *Foundations of Language*, 4(3): 225–42.

Grice, Paul. 1969. "Utterer's Meaning and Intention." *Philosophical Review*, 78(2): 147–77.

Grice, Paul. 1975. "Logic and Conversation." In *Syntax and Semantics 3: Speech acts*, edited by Peter Cole and Jerry Morgan, 41–58. New York: Academic Press.

Grice, Paul. 1978. "Further Notes on Logic and Conversation." In *Syntax and Semantics 9: Pragmatics*, edited by Peter Cole, 113–28. New York: Academic Press.

Grice, Paul. 1981. "Presupposition and Conversational Implicature." In *Radical pragmatics*, edited by Peter Cole, 183-198. New York: Academic Press.

Horn, Laurence R. 1989. *A Natural History of Negation*. Chicago: Chicago University Press.

Peregrin, Jaroslav. 2014. *Inferentialism: Why Rules Matter*. London: Palgrave Macmillan.

Sauerland, Uli. 2004. "Scalar Implicatures in Complex Sentences." *Linguistics and Philosophy*, 27(3): 367–91. https://doi.org/10.1023/B:LING.0000023378.71748.db

Soames, Scott. 1982. "How Presuppositions Are Inherited: A Solution to the Projection Problem." *Linguistic Inquiry*, 13(3): 483–545.

van Rooij, Robert, and Katrin Schulz. 2004. "Exhaustive Interpretation of Complex Sentences." *Journal of Logic, Language and Information*, 13(4): 491–519. https://doi.org/10.1007/s10849-004-2118-6

Organon F 31 (3) 2024: 217–231
https://doi.org/10.31577/orgf.2024.31302
ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

Defining Second-Order Desert

Beşir Özgür Nayır*

Received: 13 February 2024 / Accepted: 1 July 2024

Abstract: Philosophers who work on desert-adjustment within axiology often articulate the concept of desert as follows: x deserves y on the basis of z. This formulation allows for a focused examination that encompasses deservers, deservings, and desert bases. I call this *first-order desert*. This paper posits that axiology grounded solely in first-order desert fails to adequately capture our nuanced intuitions concerning desert. I contend that to construct an axiology that more effectively aligns with our desert-sensitive intuitions, we must incorporate considerations of *second-order desert*. Second-order desert is defined as follows: x deserves to live a life in which x deserves y on the basis of z. Initially, I provide a definition of first-order desert, followed by an elucidation of second-order desert. Subsequently, I explore various counter-arguments against my proposition. I defend my proposal against potential counter-arguments, demonstrating that a desert-adjusted axiological theory will be significantly better-off by incorporating second-order desert considerations.

Keywords: Desert; axiology; ethics; desert-adjustment.

* Boğaziçi University, Turkey
https://orcid.org/0000-0002-2649-2715
Department of Philosophy, Boğaziçi University 34342, Istanbul, Turkey.
ozgurnayir@gmail.com

1. Introduction

This paper posits that the existing debates on desert and desert-adjusted axiology leans heavily on what is termed as first-order desert. To prevent injustice in specific cases, I argue that second-order desert ought to be defined. This section aims to provide a concise explanation of first-order desert. The subsequent section lays out the definition of second-order desert and provides several arguments in its support. Possible counter-arguments are then introduced and assessed in the third section. Finally, the fourth section explores the potential correlation between second-order desert and luck egalitarianism.

Philosophers often articulate the concept of desert in the following manner:

Desert: x deserves y on the basis of z

Similarly, philosophers who work on desert-adjustment in axiology often contend that it is more desirable for individuals to receive what they deserve compared to being deprived of what they deserve. This notion implies that an axiology that incorporates desert-based considerations would be better-suited for capturing our desert-sensitive moral intuitions. As a result, a well-structured axiology that incorporates considerations of desert would be more preferable than a straightforward welfarist approach.[1]

For example, suppose the following statement is true in a possible world W1:

Jack deserves 1000 units of well-being.

In another possible world W2, assume that the following is true (all other things equal):

Jack does not deserve 1000 units of well-being.

Upon Jack's receipt of 1000 units of well-being in both worlds, the total well-being of each world is enhanced assuming that Jack's receipt does not entail more suffering to others.[2] An axiology that incorporates desert may reveal that W1 is more preferable than W2, as it is deemed better when an

[1] I will exclude anti-desertist arguments for the sake of this paper. See Zaitchik (1977) for a substantial defence of desertism.

[2] I will omit this possibility, as this paper focuses on a different problem.

individual receives what they deserve. A novel formula for intrinsic value may be introduced to demonstrate that W1 yields a greater expected value than W2, quantitatively. However, this paper does not delve into that aspect of the debate. It rather focuses on presenting second-order desert as a useful tool within axiology.

The comparison between W1 and W2 serves as a rudimentary demonstration of the concept behind desert adjustment, and the way we view desert is the most elementary comprehension of it. Nevertheless, the manner in which philosophers incorporate desert is typically more sophisticated. Initially, one may deserve anything on a particular basis. However, as Kagan (2014) and numerous others contend, moral desert holds greater philosophical significance. I have no intention to deny that non-moral desert may hold philosophical significance as well. However, from now on, I will be focusing on moral desert.[3] The following scenario would exemplify moral desert that is grounded in the moral worth of the agent:

> Jack deserves pleasant things since he leads a morally meritorious life.

Conversely, if Jack does not lead a morally meritorious life, one may assert:

> Jack does not deserve better things since he does not lead a morally meritorious life.

The concept underlying this interpretation of desert is straightforward. Morally speaking, if Jack contributes to the greater good in the world he inhabits – regardless of how we define what constitutes good – then he deserves better things than those who fail to contribute. For instance, if Jack assists others, actively works towards the betterment of society and humanity at large, and refrains from causing harm to others, then one may claim that his life has a greater moral value –albeit it is not the objective of this paper to establish these conditions with precision.

Philosophers have attempted various approaches to adjusting consequentialism for desert. Feldman (1995) asserts that desert increases or mitigates intrinsic value, proposing his version of desert-adjusted utilitarianism. Others, such as Gustaf Arrhenius and Bradford Skow, advocate for

[3] As otherwise anything may serve as a desert base and consequently, this debate will fail grasping our desert-sensitive intuitions within axiology.

better models of desert-adjusted axiology. (Arrhenius 2007; Skow 2012) Nearly all of the contributions to this literature pertain to what I shall define as first-order desert. In this paper, I contend that we must establish a definition of second-order desert for various reasons. One major reason behind introducing second-order desert as a useful concept will be that first-order desert does not grasp what an agent may have deserved if they were simply luckier. Another will be that if an agent's potential to flourish as a morally worthy person is not actualized due to external reasons, then it would be counter-intuitive to suggest that a desert-sensitive axiology may omit this fact. I argue that first-order desert cannot address this problem, and thus, second-order desert will be a valuable concept that can enhance a desert-adjusted axiology's intuitive appeal.

2. Second-Order Desert

I argue that the current literature exclusively focuses on first-order desert, where an individual deserves a certain outcome on the basis of a specific criterion. A desert-sensitive axiology should also consider second-order desert.

> Second-order desert[4]: x deserves to live a life in which x deserves y on the basis of z

Second-order desert pertains to the idea that an individual deserves to inhabit a world in which they get the chance to flourish as a better person, morally speaking. For example, consider Jack, whose life is defined by a certain moral worth. While he could be a better person if he had not experienced traumatic events or had access to better education, his life is characterised by these factors, culminating in a poor understanding of social and moral responsibility.

[4] I maintain this definition for the sake of simplicity. However, it actually suggests an additional desert base (let's call it w). Therefore, a more comprehensive definition would be as follows: On the basis of w, x deserves to live a life in which x deserves y on the basis of z. I appreciate Reviewer #1 for pointing this out.

I do not claim that individuals are inherently good or bad, nor that those who lead morally worthy lives are simply lucky in general. It is entirely possible that certain individuals care more than others, which may even be explained biologically by an individual's capacity for empathy. My intention is not to take a position on this matter, but rather to suggest that it is easier to accept that some individuals lead lives with less moral worth than they would otherwise have. If this is the case, then it is useful to consider second-order desert.

Assuming that there is a set of life conditions that fosters a propitious environment for an individual to lead a morally superior life, and further assuming that such conditions are present for some but not for Jack, it follows that, if *ought implies can*, and I use *can* in a weaker sense here, it is reasonable to assert that Jack cannot be judged solely because he does not lead a morally upright life, as he has not been provided with the resources necessary to flourish as a moral human being. There have been similar debates with similar motives, such as the discussion regarding responsibility as a necessary condition for something to be considered a deserving basis. (Feldman 2012; Rachels 1978; Cupit 1996) If we assume that responsibility is a necessary condition for a deserving basis, then one may claim that there is no deserving basis for Jack's past sufferings, as he is not responsible for what happened. At this point, I deem Feldman's (2012) assertion against responsibility as a necessary condition for desert bases sufficiently compelling. I believe it would be counter-intuitive to claim that one may deserve something if and only if they have at least some sort of responsibility for the deserving basis. If so, then my proposition for defining second-order desert shall remain unchallenged by such a contention. It is crucial for the sake of my argument to note that lack of responsibility may still undermine negative desert based on lack of moral worth. The following argument is based on this claim.

Now, I will make more general and intuitively appealing claims regarding morality: if morality is valuable, and leading a morally good life is consequently worthwhile, and if we exist in a world in which flourishing as a moral person is contingent on various factors, then (1) an individual may not be fully responsible for leading a morally less worthy life, and (2) a world in which everyone has the opportunity to flourish as a morally good

person is more desirable. I contend that these premises and their antecedents are accurate. Consequently, I arrive at two conclusions, both of which demonstrate the significance of second-order desert.

Regarding the first conclusion, if Jack cannot be held fully responsible for living a morally less worthy life due to a lack of conducive life conditions, then a desert-adjusted axiology cannot be just in claiming that Jack deserves less than others, as it is not completely his fault. To render justice, the axiology must compensate for Jack's second-order deserving by stating that:

> Jack deserves to live a life in which he may deserve better than his current situation in terms of his moral worth.

Regarding the second conclusion, if a world in which everyone has the opportunity to flourish morally is more valuable compared to a world in which not everyone has that chance, then axiology should prioritize the former. In the latter world, some individuals are deprived of the conditions necessary to live morally worthy lives, and to prioritize the more valuable world, axiology must consider second-order deserving by stating that:

> Jack deserves to live a life in which he has the chance to deserve better than his current situation in terms of his moral worth.

In both cases, a desert-adjusted axiology will need second-order desert.

3. Counter-arguments

One possible approach to contest this perspective is to assert that moral worth is contingent on an individual's capacity. Consider a scenario in which a person with limited financial resources is compared to the wealthiest individual in the world. In this case, the moral worth of each individual would hinge on the amount of good they accomplished in relation to their resources. If both individuals donated $10,000 to a charity, the same act of benevolence would hold varying degrees of significance for their moral worth. For the middle-class person, it would be a momentous feat of goodness, but for the world's richest individual, it would not be as substantial. Consequently, acts of goodness are not absolute, but instead, they have a

marginal value. If so, then we may address the problem without referring to second-order desert.

This counter-argument is valid to some extent, particularly when it pertains to quantifiable means of benevolence such as charity. However, it falls short when we consider the full range of actions that impact moral worth. Certain acts of goodness may not be readily quantifiable or susceptible to compensation with money, such as showing kindness and compassion to others. In such cases, we must establish a means of measuring the difference in the marginal value of one person's kindness compared to another's. Unfortunately, it is difficult to find a plausible method to defend this counter-argument.

An alternative perspective contends that a desert-adjusted axiology does not necessarily diminish the amount of well-being received by agents with less moral worth. Instead, such an axiology would merely indicate that it is preferable for individuals to obtain what they deserve. This implies that those with neutral or even negative desert values may not receive less than they would within straightforward welfarism. Some philosophers who examine desert-adjustment, such as Skow (2012), emphasise that their work excludes negative desert, as it requires a more sophisticated account to argue that individuals with negative desert values should receive less well-being, or even lose some well-being –unlike Feldman, who explicitly states that negative desert mitigates the intrinsic value of pleasure. (Feldman 1995) Nonetheless, suggesting that a person with a low moral worth should be penalised by a reduction in their well-being is more costy and may lead to intuitively challenging results. Consequently, if we avoid making such a claim, individuals who cannot flourish in terms of moral worth will not lose anything.

Despite agreeing that considering negative desert values leads to difficult debates, I do not believe this counter-argument is tenable. The necessity of accounting for second-order deserving is not merely because those who cannot flourish will suffer from a lack of moral worth. Rather, it is necessary to maintain fairness within a desert-sensitive framework. When Jack is unable to flourish as a morally good person, he may be receiving less than what he deserves.

	W3	W4
A1	Deserves 1000, gets 1000	Deserves 1000, gets 1000
A2	Deserves 1200, gets 1000	Deserves 1200, gets 1200

Table 1

Let us assume that in the table above, W3 is the possible world where Jack is unable to flourish as a morally good person, W4 is the possible world where (all else being equal) Jack can flourish, A1 is the version of Jack who does not prefer living a morally worthier life even with the chance to do so, and A2 is the version of Jack who would choose to live a morally worthier life if given the opportunity.

If Jack is A1, then we cannot argue that A1 deserves to live in W4 any more than he deserves to live in W3, as there will be no difference in terms of expected value. However, if Jack is A2, then a desert-adjusted axiology should prioritise W4 as it offers greater expected value. Therefore, prioritising W4 has potential benefits overall, let alone the fact that it responds better to our desert-sensitive intuitions.

It would be fairly implausible to assert that there would be no cost associated with prioritizing W4. If the cost of prioritizing W4 is rationally expected to exceed the potential benefits, then we may not be morally justified to prioritize it. However, this is unlikely to be the case unless one adheres to strict welfarism. For a welfarist, prioritizing and pursuing W4 could diminish the overall utility when compared to other possible worlds where overall utility is greater. But from within a desert-sensitive framework, this would only be the case if, in W4, Jack gets the chance to flourish at the expense of others who lose their opportunity to flourish or are deprived of their deserved well-being. This would reduce the overall expected value of that possible world. Since Jack represents anyone who did not get the chance to flourish, I do not anticipate any such problems. However, even then, it would not make the concept of second-order desert less useful.

One may also argue that when considering desert-adjustment, it is better to take a whole life approach. This would mean that we should evaluate a person's entire life in order to determine what they deserve overall.

According to this approach, "[...] time drops out from further consideration: we look at lives as a whole, to see what one deserves (overall), and whether one has received it (overall)." (Kagan 2014, 11) If so, then not being able to flourish as a better moral agent may not be significant, as we would compare the desert value of a whole life with the received value. This approach would already acknowledge the lack of receipt in the past (such as not growing up in a peaceful environment). This way, it would recognize certain facts that make the agent incapable of flourishing as a better moral agent without invoking second-order desert.

Let us assume that the whole life approach holds up well. If it does, a charitable interpretation of it would recognise the existence of certain undeserved states of ill-being that eventually impose limitations on future actions of the agent. Such a circumstance may result in a life with less moral worth than the one the agent would otherwise have had. However, even under these ideal conditions, I contend that the concept of second-order desert is a superior tool for the reason I explain below. To illustrate this, I present a thought experiment that exemplifies a scenario in which the concept of second-order desert does a better job explaining the situation compared to the whole life approach. This thought experiment was also formulated by Brad Hooker[5]:

Suppose there are three factories situated in close proximity to a river, and let us further suppose that the river will become polluted if more than a third of the waste produced by these factories is discharged into it. Consequently, in order to avoid polluting the river, at least two of the three factories must safely dispose of their waste through methods that do not involve dumping it in the river. Let this method be safely burning the waste. However, the cost of burning the waste is significantly higher than simply dumping it into the river. Furthermore, once the river has already been polluted, any individual factory's decision not to dump its waste into the river will have no significant impact on the overall pollution levels. In other words, choosing not to dump waste into a polluted river does not provide any tangible benefits.

[5] This is also where Hooker grasps an intuition similar to the problem I will show in the thought experiment. The difference is that my version focuses on an involuntary loss of potential moral worth. See Hooker (2002, 124-5).

Scenarios	A		B		C		D	
	Discharges	Burns	Discharges	Burns	Discharges	Burns	Discharges	Burns
Factory 1	x		x			x		x
Factory 2	x		x			x		x
Factory 3	x			x	x			x
Outcome	Polluted		Polluted + Costy		Clean		Clean + Costy	

Table 2

This table offers a sufficient number of possible combinations for our purposes. Let us assume that Factory 3 is unwilling to discharge its waste. Comparing scenarios A and B, when the other factories do not choose to go green, it does not seem rational for Factory 3 to bear the cost of burning its waste. In this case, scenario A appears to bring more overall well-being compared to scenario B. Likewise, comparing scenarios C and D, when the other factories choose to go green, it does not appear rational for Factory 3 to incur the cost of burning its waste. In this comparison, scenario C appears to bring more overall well-being compared to scenario D.

In both cases, it is necessary for Factory 3 to discharge its waste into the river in order to increase overall well-being. In scenario C, factories 1 and 2 spend more resources to protect the environment, while Factory 3 manages to evade this responsibility, even though the owners of Factory 3 were willing to make the same sacrifice for the environment. Considering the owners of these three factories, does this mean that the owners of Factory 3 lead morally less worthy lives compared to the owners of the other factories? After all, by mere luck, they acted less environmentally responsible compared to the others. If the answer is no, then how shall we recognize the sacrifices of other factories and eventually praise them? If the answer is yes, then how exactly shall we justify the lack of moral worth in the lives of the owners of Factory 3?

Arguing that the owners of the third factory live a morally inferior life due to their waste management strategies, when all else is equal, seems implausible. It is equally implausible to suggest that the owners of the first two factories live equally worthy lives, even though they made a sacrifice that the owners of the third factory did not. Using only first-order desert, we may suggest that the owners of the first two factories deserve better

things, as they sacrificed more for the sake of a better world. Second-order desert is a more useful tool when recognizing the moral worth of the owners of the third factory: they deserve the opportunity to flourish as morally better agents, as they intended to do the right thing.

Assuming that all three factories aim to adopt environmentally friendly practices in scenario C, the third factory was denied the opportunity to flourish. The same situation emerges in scenario A, where only the third factory wishes to incinerate its waste but could not do so in order to increase overall well-being. In scenario C, in terms of the distribution of well-being, one can claim that (1) the owners of the first two factories have first-order desert because they contributed more to the world, and (2) the owners of the third factory have second-order desert because they would have contributed more to the world if they had the opportunity. Similarly, in scenario A, one can argue that (1) the owners of the first two factories lack first-order desert since they failed to contribute to the world when they could have, and (2) the owners of the third factory possess second-order desert because they would have made a greater contribution if given the chance.

What I've been describing as second-order desert might actually be better understood as a lack of opportunities, which, for some reason, elude moral agents. The concept that one can deserve opportunities isn't a new one (Schmidtz 2006; Feldman 2016). Considering the central thesis of this paper, one could argue that defining second-order desert is unnecessary, as it can be articulated in first-order terms as follows:

> Deserving opportunities: x deserves the opportunity to get y on the ground z.

I recognize that the first-order formulation mentioned earlier will encompass certain scenarios that proponents argue should fall under the concept of second-order desert. For instance, a person rightfully deserves the opportunity to thrive as a morally virtuous individual. In this context, you might wonder what sets my proposal apart. I have two responses to this critique.

First and foremost, the concept of a person deserving an opportunity is inherently forward-looking. For instance, consider Jack, who deserved a better education or upbringing 20 years ago but didn't receive it. Looking at this from today's perspective, it indeed makes sense to view Jack as

someone who deserved those opportunities in the past. However, if we adhere to the first-order formulation, rectifying the inequalities Jack endured would require us to somehow provide him with the opportunities he missed. In essence, first-order desert designates a yet-unfulfilled receipt. On the other hand, second-order desert serves as a placeholder that captures past inequalities affecting Jack's current moral worth. It may not seem intuitively plausible to claim that Jack deserves the opportunity to attend a good primary school now, as we can't turn back time. Nevertheless, it does make sense to assert that Jack deserves to have deserved such opportunities, even though some past inequalities are nearly irreversible. The second-order formulation doesn't necessitate us to offer Jack what he deserved in the past; instead, it prompts us to consider a certain well-being that he currently lacks. In essence, rather than dwelling on the missed opportunities, it focuses on a specific desert value that arises from the absence of such opportunities.

A second response to this objection, somewhat intertwined with the first, is that second-order desert widens the scope of what one might have lacked in the past. Some things are not mere opportunities but rather fundamental. Growing up in a mediocre household, for example, is scarcely perceived as an opportunity, yet it becomes challenging to argue that a child did not deserve it if they lacked it. Similarly, attending primary school, while not strictly an opportunity, is more of a foundational aspect of life. Even though many children still lack this privilege, it's regarded as something more fundamental than a mere opportunity. Given that second-order desert aims to encompass a certain sense of desert value arising from either inequalities or simple (mis)fortune, I contend that it finds greater utility within axiology.

The final objection to my proposal that merits consideration can be somewhat intricate. While I advocate for the incorporation of second-order desert, one could argue that to better capture our desert-sensitive intuitions, we need to define third-order desert, and this might lead to claims for fourth-order desert, and so on. While this may apply in specific situations, the utility of defining third- or fourth- (or n-th-) order desert appears questionable. The distinction between first-order and second-order desert is akin to the distinction between desertist axiology and non-desertist axiology, as it fundamentally alters our perspective. However, the difference between

second-order desert and third-order desert doesn't seem as pronounced, given that the primary purpose of second-order desert is to grapple with certain inequalities that are otherwise challenging to address.

Nonetheless, a compelling critique could present a plausible thought experiment necessitating the definition of third-order desert. In such a case, second-order desert would still retain its justification as a valuable (although not exhaustive) tool. Personally, I wouldn't find this problematic.

4. Second-order desert and luck egalitarianism

The intuition behind the concept of second-order desert bears similarity to the motivation behind luck egalitarianism. Luck egalitarianism seeks to address the injustice that arises from involuntary differences between individuals leading to inequality. (Rawls 2020; Dworkin 2000; Arneson 2018; Cohen 1989) This is similar to how second-order desert recognises certain inequalities resulting from causes outside an agent's control. This similarity can be interpreted in two ways. It may suggest that second-order desert is a luck egalitarian tool. Alternatively, it can be seen as an independent concept that complements luck egalitarianism.

Firstly, one may contend that second-order desert is unnecessary and we should simply embrace luck egalitarianism. However, this approach may not suffice as luck egalitarianism is solely concerned with theories of distributive justice and does not contribute to axiology in the way the concept of second-order desert does. When incorporated within a desert-adjusted axiological theory, second-order desert can further improve consequentialist theories as first-order desert did when Feldman first presented his desert-adjusted utilitarianism in response to Rawls' critique of utilitarianism. (Feldman 1995) Thus, despite the success of luck egalitarianism in capturing our desert-sensitive intuitions, second-order desert still promises a substantive contribution.

Secondly, one may view second-order desert as a means of linking consequentialist axiology with luck egalitarianism. A well-constructed desert-adjusted axiology can provide a sound theoretical foundation for luck egalitarianism and respond to philosophical questions raised about it. By showing how inequalities resulting from luck lead to an inferior moral world, a

desert-adjusted axiology can demonstrate why such inequalities are undesirable.

5. Conclusion

The commensurability of first-order desert and second-order desert remains a contentious issue, and even if they are commensurable, a plausible method for doing so needs to be established. These questions are left for future research. The aim of this paper was to introduce the concept of second-order desert as a valuable tool. The prevalent notion of desert in existing literature only utilises first-order desert. However, in certain morally significant circumstances, we need to take into account second-order desert. The concept of second-order desert could aid us in identifying different types of inequalities in our present world within a desert-adjusted axiology. Additionally, it could assist us in comprehending the role of contingent factors in determining moral worth and in making it more convincing regarding moral desert. Further exploration will demonstrate the usefulness of the concept of second-order desert. Nonetheless, it is an idea that undoubtedly warrants more attention.

Acknowledgements

I would like to thank Nurbay Irmak and Hüseyin Kuyumcuoğlu for their valuable comments and contributions.

References

Arrhenius, Gustaf. 2007. "Meritarian axiologies and distributive justice." *Hommagea Wlodek: Philosophical papers dedicated to Wlodek Rabinowicz* 1-43.

Arneson, Richard J. 2018. "Equality and Equal Opportunity for Welfare." In *The Notion of Equality*, edited by Mane Hajdin, 237–53. London and New York: Routledge. https://doi.org/10.4324/9781315199795

Cohen, Gerald A. 1989. "On the Currency of Egalitarian Justice." *Ethics*, 99(4): 906–44. https://doi.org/10.1086/293126

Cupit, Geoffrey. 1996. "Desert and Responsibility." *Canadian Journal of Philosophy*, 26(1): 83–99. https://doi.org/10.1080/00455091.1996.10717445

Dworkin, Ronald. 2000. *Sovereign Virtue: The Theory and Practice of Equality*. Cambridge, Massachusetts: Harvard University Press. https://doi.org/10.2307/j.ctv1c3pd0r

Feldman, Fred. 1995. "Adjusting Utility for Justice: A Consequentialist Reply to the Objection from Justice." *Philosophy and Phenomenological Research*, 55(3): 567–85. https://doi.org/10.2307/2108439

Feldman, Fred. 2012. *Utilitarianism, Hedonism, and Desert: Essays in Moral Philosophy*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139174978

Feldman, Fred. 2016. *Distributive Justice: Getting What We Deserve from Our Country*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198782988.001.0001

Hooker, Brad. 2002. *Ideal Code, Real World: A Rule-Consequentialist Theory of Morality*. Oxford: Oxford University Press. https://doi.org/10.1093/0199256578.001.0001

Kagan, Shelly. 2014. *The Geometry of Desert*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199895595.001.0001

Rachels, James. 1978. "What People Deserve." In *Justice and Economic Distribution*, edited by John Arthur and William H. Shaw, 167–96. Hoboken, New Jersey: Prentice-Hall.

Rawls, John. 2020. *A Theory of Justice*. Revised edition. Cambridge, Massachusetts: Harvard University Press. https://doi.org/10.2307/j.ctvkjb25m

Schmidtz, David. 2006. *The Elements of Justice*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511817519

Skow, Bradford. 2012. "How to Adjust Utility for Desert." *Australasian Journal of Philosophy*, 90(2): 235–57. https://doi.org/10.1080/00048402.2011.572079

Zaitchik, Alan. 1977. "On Deserving to Deserve." *Philosophy & Public Affairs*, 370–88.

Organon F 31 (3) 2024: 232–277
https://doi.org/10.31577/orgf.2024.31303
ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

Multiple Realizability and Disjunction for the Special Sciences

Roque Molina Marchese*

Received: 16 November 2023 / Accepted: 8 July 2024

Abstract: One way to secure the autonomy of special sciences like psychology is to block reductive strategies by assuming that higher-order properties in psychology are multiply realizable. Multiple realizability would then secure both metaphysical irreducibility and dependency by exploring the variety of ways in which higher-order phenomena can be realized in different systems. Originally, a promising way to understand this variability was in terms of the possible realization role played by property disjunction. However, the non-projectability of disjunctive predicates into explanatory generalizations undermines the multiple realizability strategy mainly because a condition for these generalizations to have scientific weight is that they be based on the existence of natural kinds. Traditionally, disjunctive properties have no reference to kinds. In this paper I explore the character of disjunctive properties as cases of homeostatic property clusters sufficient to be classified as genuine natural kinds, and the consequences for the question of the autonomy of the special sciences.

* Charles University and University of Skövde
https://orcid.org/0009-0006-7304-492X
Department of Philosophy and Religious Studies, Charles University, nám. Jana Palacha 2, 116 38 Prague 1, Czech Republic
Department of Cognitive Neuroscience and Philosophy, University of Skövde, Högskolevägen, Box 408, 541 28 Skövde, Sweden
roque.molina.marchese@his.se

Keywords: Disjunctive properties; homeostatic property clusters; multiple realizability; natural kinds; realization; special sciences.

1. Introduction

Since its introduction into the philosophy of mind (Fodor, 1997; Putnam, 1992), multiple realizability has focused on the conceptual relation between non-reducible properties and their physical realization by means of disjunctive property bases. The idea was to block type-to-type identities that were thought to be necessary to achieve the reduction of one property to another (Smart, 1959). The blocking strategy consists in combining the realization of higher-order properties with a *variability* in the ways in which that realization can be achieved. This variability can be thought of as sets of disjunctive properties satisfying a one-to-many realization relation between a target phenomenon and its explananda. Disjunction of properties amounts to the idea of physical components and their properties forming groups, the many part of the realization relation, held together through Boolean "OR" logical operator. The members of the group function as realization bases for other properties and most importantly membership is conditioned by variability such that one or more members can be replaced or added without that affecting their realization potential. By making use of sets of disjunctive properties, the proponents of multiple realizability want to make irreducibility and ontological dependence compatible with each other; higher-order properties like mental ones are ontologically based on their physical bases, but they are nonetheless properties of a different kind, which do not reduce to the latter (Clapp, 2001; Jaworski, 2002). The irreducibility of mental properties is then achieved through the kind of variability provided by disjunction which then is thought to effectively block type to type identities. That kind of variability in realization bases given by disjunction of properties represents then the multiple realizability character of higher-order properties. As a result, multiple realizability represents, metaphysically speaking, a case of non-reductive physicalism. Critics hold that multiple realizability is a flawed metaphysical project given that what first seemed to be a virtue in the use of disjunctive realizers – i.e., the use of disjunction to establish irreducibility – actually undermines the

coherence of the multiple realizability thesis itself (Dosanjh, 2021; Francescotti, 1997). Therefore, according to the critiques, disjunctive properties make irreducibility prima facie possible, but at the high cost of making multiply realized properties scientifically wild or unprojectable into explanatory predicates. That is, using disjunctive properties as the bases for establishing explanatory relations between physical bases and higher-order properties renders whatever predicates could be derived from those relations explanatory empty. Disjunctive properties cannot figure in explanatory frameworks, because explanatory generalizations cannot be based on properties whose inherent variability does not represent cases of *stable* natural kinds. Moreover, to the extent that multiply realized properties are part of the conceptual and explanatory frameworks of special science, this criticism would have implications for the integrity of those sciences as well (Bechtel and McCauley, 1999).

2. Disjunction of properties

A set of properties is disjunctive with respect to a higher-order phenomenon, such as a certain type of memory, if the properties in that set either play the defining functional role characteristic of it, or the set represents the actual physical constituents that make up the memory as such (Tonegawa et al., 2015). Further, the members of the set *vary* from time to time and so cannot be the same each time they happen to instantiate the same kind of higher-order phenomenon. Disjunction then means that the members of the realizing set can vary, and yet the memory remains the same by keeping properties presenting its functional and phenomenological integrity stable. Thus, we could illustrate the disjunctive relationship between e.g. a cognitive function such as episodic memory and the brain states responsible for its realization by means of the software/hardware metaphor for the mind-body relationship. For example, the memory of my tour of the Palazzo Pitti is disjunctively realized whenever different physical brain bases computationally implement the occupant role defining that memory, or the very presence of those physical brain bases materially constitutes the presence of that episodic memory (Quiroga, 2020, 2023; Suthana et al., 2021). The software here, i.e. the episodic memory of the tour, is realized by a disjunction

of brain states and their properties, i.e. the hardware. Importantly, episodic memory in general is regarded as a intelligible psychological phenomenon whose physical dependency yet cognitive autonomy from its brain bases is secured by the fact that the brain bases involved in the realization of the memory would vary from occasion to occasion either as functional role occupants or as physical constituents (Anderson, 2016; Colaço, 2022; Edelman and Gally, 2001; McCaffrey, 2023). Moreover, the character of this variability represents physical states extending both within and across individuals of the same and different species (Noppeney et al., 2004; Noppeney et al., 2006). In other words, realizing the same general property by disjunction involves individuals of different species and individuals within the same species (Minelli, 2019). Therefore, in the case of the memory of Palazzo Pitti, we could say that the identification of the set of disjunctive properties responsible for this higher-level phenomenon is tantamount to the identification of the physical causes responsible for its realization. Moreover, based on the disjunctive character of the set, an explanatory relation between the set and the phenomenon can be developed.

3. The dilemma of disjunction

So far, so good, but opponents point out that the problem with disjunctive properties featuring in explanatory models is that they are not commonly classed as natural kinds, and thus do not lend themselves to projection into theoretical generalizations (Antony, 2003). A common expectation and implicit requirement for the development of scientific explanations is that the models involved in these explanations range over the existence of certain properties that qualify as natural kinds (Boyd, 2021). Whatever the understanding of what it means to be a natural kind, a basic condition is that this kind be relevant to the development of testable hypotheses and theories within the framework of established scientific practices. Therefore, if disjunctive properties *do not* fall under the natural kind category then they cannot constitute nor contribute to the formation of explanatory generalizations. In order to respond to this objection some important observations regarding the relationship between disjunction and realization are in place here. Thus, disjunctive properties can realize other properties through

the implementation of functional roles, as well as through their constituents/component characteristics. Thus, they cover both ideas behind *physical* realization: realization as a constitutive physical relation and realization as occupant causal role implementation or what even amounts to the idea of psychophysical functional realization. Within the framework of associating mental properties with brain states, disjunction suggests that realization can be ascribed to material constitution, psychophysical functional realization, or potentially a combination of the two. Sets of disjunctive properties could even correspond to *restricted* instances or partitions of physical properties realizing the functions responsible for higher-order properties. From a material point of view, the occurrence of disjunctive properties occupying restricted partitions of space and time is the necessary basis on which higher-order phenomena like memory and mental state contents can be materially grounded. Nonetheless, not all partitions of a set of disjunctive properties would count as necessary for the realization of the higher-order phenomenon. Only those partitions that currently satisfy a sufficient condition for entailing realization would count; that condition being that the disjunctive bases instantiate the *right* type of constituent or the *right* type of functional role. Nevertheless, it is crucial for the character of disjunction that the physical states being members of the sets sufficient for the realization of the higher-order phenomena *vary* and fluctuate from time to time. In other words, disjunctive physical states must be in principle interchangeable as members of realizing sets. Therefore, disjunctive properties participating in the realization relation denote a *relative* condition not an absolute one. The combination of a limited number of set members with the requirement of variation represent key aspects behind the relative condition characteristic in the use of disjunction as realization bases; the limitation ensures that a relevant combination of properties is sufficient and necessary for the realization to happen and the variation ensures that the realization relation is both robust and flexible. Yet, critics want to point out that the relative character behind the choice of disjunctive physical states is in fact an *arbitrary* decision made by the observer. The question of which properties are relevant set members and how variability should be measured track at best an epistemological condition dependent on the observer but does not necessary correspond to an ontological condition independent of the observer

(Bechtel and Mundale, 1999). In the range of possible realizers available through disjunction it is the observer who picks up one or many different sets that could play the role of realization. The observer's decision could be based on implicit pragmatic and methodological issues. Thus, the critics argue that the character of the ontology is guided by the observer's epistemology. A consequence of this is that if the kind of variability offered by disjunction is based on arbitrary decisions then that would inevitably limit the generalizability of possible explanations based on disjunctive variability. Thus, if organisms from species O and O* are in physiological state H sharing the common experience *M* of *feeling hungry* at time *t*, then H is realized disjunctively in O and O* at *t* by a set **S** of relevant yet different base properties responsible for H in both O and O*. Furthermore, the property members of **S** will not be the same at another time *t** the higher-order property ***M*** is being instantiated in both O and O* as a function of H. Moreover, not only do the members of **S** vary from time to time but so even the very sets themselves. In other words, the realization relation between M and set S is not a necessary relationship and cannot be projectable to new cases. This situation does not support *generalizations* because realization by disjunction is a relative and *contingent* relation. Given the broad range of variability typical for disjunctive properties and their corresponding instantiating physical states, any attempt to make use of them in explanatory propositions is doomed to fail. The domain of possible explananda becomes *too broad* and irregular because of the very range of variability and this is which makes disjunctive properties *explanatory wild*. Thus, the critics conclude that given such an irregular character, disjunctive properties when used as examples of multiple realization actually undermine any possible explanatory power behind the thesis of multiple realizability (Sober, 1999). Therefore, the multiple realization of higher-order properties on the basis of disjunction is not an option that is either theoretically or scientifically sound (Schneider, 2012). The sort of contingency involved in the use of disjunctive properties affecting explanatory generalizability is symptomatic of disjunctive properties being *explanatorily lawless* i.e., they do not pick up nomological relationships because such properties do not classify as instances of natural kinds. Thus, disjunctive properties as realizers cannot be part of scientific explanations if an essential condition for the properties to

be part of those explanations is that they enter into lawful regularities by means of picking up natural kinds. Since only natural kinds represent the kind of general and recurrent entities that make it possible for them to appear in explanatory models, disjunctive properties that lack lawfulness and generalizability disqualify them as examples of natural kinds. So, on the one hand making use of disjunction for realization supports the irreducibility emphasized by the multiple realizability thesis; disjunctive properties provide for the kind of variability required for the multiple one-to-many character of the realizing bases involved. On the other hand, disjunctive properties not being examples of natural kinds means the realization relation between $\boldsymbol{M}$ and $\mathbf{S}$ should be understood as a *brute* fact. This situation creates a conflict for the proponent of multiple realizability: disjunction seems to secure non-reducibility at the cost of the realization relation being a brute fact and so apparently it becomes non-tractable for scientific analysis and research.

As an answer, the proponent of multiple realizability would like to emphasize the following observation. If our best scientific evidence is indicative of mental properties in general being realized by brain states and if those states correspond to correlated sets of disjunctive properties, then because of disjunction, or better expressed in virtue of being disjunctive, we would lack the *bridge laws* required for establishing the conditions necessary for the reduction of M to $\mathbf{S}$. Bridge laws provide the theoretical matching between the properties of one level of explanation in terms of properties pertaining to another level of explanation (such as in the translation of temperature as a phenomenon into the language of statistical thermodynamics). Once the bridge laws are in place then the possibility of reduction between theories and apparently different properties becomes a serious possibility (Hempel, 1988). Now, disjunctive properties do not apparently meet the requirement for entering into lawful explanatory frameworks, not so either into the frameworks of bridge laws. Thus, to the extent that the properties of a supervenient domain M are identified with the properties of its subvenient physical domain P, disjunctive properties as physical realizers of M actually block the reduction of M to P because disjunctive properties in general do not fall into bridge laws. There is further an interesting conceptual similarity here between the mind being *anomalous* and disjunctive

properties being *contingent* in general. The contingent and non-generalizable character of disjunctive properties reflects in the explanatory anomalous character of the mind (Davidson, 2006). The common denominator for both the contingency of disjunction and the anomalous character of the mind is that both exempt of falling under the scope of bridge laws and whose absence prevents reduction. Thus, in the context of non-reductive physicalism and using the computer metaphor of mind for means of illustration, the mind would be metaphysically determined by the physical nature of the brain, yet the mind, at least at the algorithmic functional level, operates in ways undetermined by the purely physical processes of brain i.e., the hardware level. Disjunctive properties as physical realizers metaphysically determine other properties but are themselves explanatory undetermined because they are contingently based. Thus, such properties realize other properties but are exempted from falling into bridge laws as a precondition for establishing theoretical reduction among properties belonging to different levels of organization. This should be good news for the *non-reductive* physicalist appealing to multiple realizability. Nonetheless, this condition leaves us with the uncomfortable situation of the relation between $\boldsymbol{M}$ and **S** being contingent and not strict, and so again metaphysically speaking, wild. The sciences dealing with properties like *M* and H, special sciences like psychology and physiology, would under such conditions of realization be context-dependent, and the inferred regularities invoked in their explanations would lack explanatory and metaphysical depth in the strict sense of the fundamental sciences like physics. Furthermore, an additionally uncomfortable effect against the supporter of non-reductive physicalism appealing to multiple realization is that if the relation between $\boldsymbol{M}$ (higher-order) and **S** (realizing bases) represents a non-necessary condition, it implies the possibility at least in principle of **S** not being physical at all. Disjunctive realization being a brute fact means that whatever alternative sets of disjunctive properties suitably arranged would instantiate the function and constitution of the higher-level phenomenon. It is not metaphysically speaking a necessary condition for those properties to be *physical* at all; even non-physical states and their properties would do the job and this is bad news for the non-reductive *physicalist*.

Thus, supporters of multiple realizability end up in a dilemma: on the one hand, making use of disjunctive properties we allow for irreducibility at the expense of explanatory powers and generalization; on the other hand, we can opt for giving up disjunction and hold to generalization to secure explanation, but then we sacrifice the variation provided by disjunction and jeopardize irreducibility. Thus, disjunction and variability of realizing properties as required by multiple realizability go hand in hand. As a result, when variability as a key component to multiple realizability becomes jeopardized by rejecting disjunction the very idea of irreducibility is rendered hollow. Even more, when appealing to disjunction as a basis to avoid reduction the explanatory and ontological status of the special sciences like Psychology seems also jeopardized. As we see, many things stand and fall with the consistency of multiple realizability and the use of disjunction as a means of variability to secure irreducibility. To come out of this dilemma we need to deal with some specific issues to provide an answer to the critics: Could we eliminate one horn of the dilemma by securing the condition of disjunctive properties being natural kinds? Can we keep irreducibility by means of disjunction and still secure explanatory powers for special sciences? Does variability or multiplicity of realization require disjunction? Key to approaching the dilemma is what we understand by a *natural kind*.

4. Disjunctive properties as natural kinds

On an essentialist understanding of natural kinds, a property G is a natural kind because it is an indispensable and constitutive element of reality as identified by our best current scientific practices (Devitt, 2021; Wilkins, 2013). Property G is then included in the ontological furniture of the world. Property G is a natural kind also in the sense that it has the power to reliably contribute to our scientific understanding whenever it is included in our best theoretical efforts to describe the world. G is therefore also an essential part of our epistemic systems and models; it enters into explanatory predicates mainly in terms of the effective differences it makes in the causal network of the world (Khalidi, 2018). Most importantly from the essentialist point of view a natural kind is independent for its existence from human reality; where there no human cognizers there will still be natural

kinds such as elementary particles. The ontological and epistemological conditions that characterize essential kinds are also reinforced by their applied and pragmatic significance, such as whenever our best technical achievements derive their success from the existence and character of these kinds in the background of reality. Thus, advanced medical equipment such as magnetic resonance spectroscopy would not provide successful diagnostic aid if *protons* were not cases of a natural kind basic to the constitution and function of the measuring device (Soares and Law, 2009). Interestingly, from the point of view of special sciences like psychology the question arises: would theoretical thinking about the workings of the mind provide successful guidance and understanding at all if mental structures such as *schemas* and *mental representations* were not also cases of higher-order natural kinds fundamental to the constitution and function of the mind? In other words, are mental and cognitive structures also cases of natural kinds on pair with protons and electrons such that they constitute kinds of their own different in character from those of the basic sciences but nonetheless as real as protons and electrons? It is also of paramount importance for the essential character of natural kinds that they take part in and play a role in lawful regularities. According to these defining standards, disjunctive properties do not capture the conditions for instantiating essential kinds. Now, not any natural kind is an *essential* kind. The scope of essential kinds could be limited to the microstructure of the world. Still, it is an open possibility that other levels of reality and the sciences operating on them make use of kinds whose character and behavior differs from the ones used in the basic sciences. Thus, it can be observed that many of the kinds that sciences such as biology and psychology deal with in categorizing and classifying their subject phenomena are different from those in sciences such as chemistry and physics (Cacioppo and Tassinary, 1990). As an example, the concept of *biological species* is basic to biological theorizing but recognizably heterogenous enough not to count as an *essential* natural kind in the straightforward sense of basic kinds such as electrons (Ereshefsky, 1992). Instead of an essentialist view of natural kinds, fields like psychology and systems biology make use of a pluralist view on kinds and understand them as clusters of homeostatic properties; *homeostatic property clusters*, HPCs (Boyd, 1991, 2021; Bruggeman and Westerhoff, 2007). Here, a cluster of properties

shows a *recurrent pattern of explanatory stability* whenever used to explain a complex phenomenon. HPCs are *stable enough* to provide epistemic guidance in the construction of explanatory models, yet *flexible enough* to reflect the context dependent and often malleable ontology of phenomena studied in psychology and biology (Onishi and Serpico, 2022). Thus, HPCs are not the sort of basic/atomistic objects like elementary particles and their properties. Rather, they reflect the more collective and contextual character of the interdependent qualities found in higher-order phenomena and their defining properties (Magnus, 2014). What is important for understanding HPCs as cases of natural kinds is their ability to function as heterogeneous interconnected clusters of properties capable of representing conditions in the causal structure of the world, or at least that part of the world under investigation (Wilson et al., 2007). Making use of HPCs as examples of natural kinds points to the *plurality* of relevant but interdependent properties that when causally combined aid to the metaphysical and scientific understanding of the phenomenon in question (Esfeld, 2005). Thus, in the economics of psychological phenomena, having a desire for water is not an isolated property but rather a conglomerate, a cluster of adjacent and relevantly interconnected physiological conditions, mental states and behaviors, like sensations, intentions and actions all contributing to the formation and identification of the desire. My need for water is then at the psychological level an example of the mental kind *Desire* which represents a contextual and distributed set of properties including among other memories, beliefs and decisions. As a kind then such contextual properties act together as *one* causally coherent set endowed with explanatory powers relevant for the type of behavior being exposed; my desire explains my behavior. The causal coherency of such properties reflects in the homeostatic character of the clusters representing their level of recurrence and stability. Correspondingly, at the neuronal level of implementation, it is not unreasonable to assume that interconnected assemblies of neurons implementing integration and segregation processing, supported by neuroplastic mechanisms, would reflect the contextual and distributed character of mental phenomena such as beliefs and desires (Seitz and Angel, 2012). Now, referring to collective and contextual clusters of properties seem counterintuitive and do not reflect our intuitions of what we mean by *a* natural kind being an *essential* and

unique property integral to nature´s furniture. Yet, the more interdependent and contextually sensitive a phenomenon is such as memory consolidation and phenotypic adaptation, the more different yet relevant clusters of properties act *together* to ground the very phenomenon (Lemeire, 2021).

The boundaries of the clusters at the subvenient constituent level are permeable enough to allow one and the same constituent property to vary as a member and participate in the constitutive function of other clusters without endangering the overall stability of the higher-level phenomenon being realized (Brigandt, 2003; Dewhurst and Isaac, 2023). Such clusters are robust and flexible enough to reflect the resiliency of the phenomenon itself (Austin, 2020; Chirimuuta, 2018). For example, biological species is not a term that refers to a monolithic, unchanging kind, but rather a concept that denotes an inherently variable category depending on its explanatory use in different fields of biological investigation (Barberousse et al., 2020). Recalling one part of our dilemma: the status of disjunctive properties as natural kinds, the idea then is to approach a solution by way of thinking of disjunctive properties *as* cases of HPCs. The realization of a mental property M by a set of disjunctive properties P must reflect, at the level of the realizers, the flexibility and permeability necessary to match the contextual and interdependent character of the property M being realized (Balari and Lorenzo, 2019). In a sense then, the degree and level of granularity between M and P matches both explanatory and componentially. Thus, the kind of *explanations* framed in the vocabulary of a science Ψ acknowledging of the role of P bases, reflect aspects of contextuality and conceptual interdependencies present in properties both across P and M-levels. Hence, in terms of property instantiation; the sort of component properties at the realization level P show interrelatedness and malleability in a way the same as the properties at the level of M. So, if mental property M* represents the single *belief* "I'm thirsty now", then according to the idea of homeostatic property clusters, M* is a case of a *composite* natural property belonging to the broader categorical type ***Belief*** to which cases of M* belongs. Here M* represents a case of the natural kind ***Belief***. Further, M* is a composite because there are many factors contributing to the instantiation of M* on both the psychological and physiological levels. Thus, my single belief that M*, is a *plural* construct tracking the existence of relevant

causal networks stable enough to keep its meaning and significance as a constitutive and explanatory property. In other words, M* plays an explanatory role in the current psychological state that I happen to be in, because it traces causal roles and correlations that explain my behavior. Importantly as well, the set of realizing properties P under which M* subsumes is *also* a composite of properties; a disjunctive set of physical states and their properties that are causally relevant for the instantiation of M* and still homeostatic or stable enough to guarantee the ontological status of M*. Therefore, if disjunctive properties correspond to HPCs and homeostatic property clusters classify as members in the ontological furniture of the world, then we may have a way to secure the status of disjunctive properties *as* natural kinds both at the psychological and at the physical level of realization. To repeat, the way disjunctive properties realize a second order property is either by means of functional role implementation or by material constitution. Nothing seems to prevent disjunctive realizing properties from taking on the defining character of homeostatic property clusters, so long as they make an identifiable causal contribution to the relations they enter into the world. What matters is that the clusters are stable, homeostatic, but permeable enough to implement functional role descriptions and to enter into material constitutive relations characterized by variability and plurality. Now, the critics may rebut that basically, this is the very requirement that enables even opponents to multiple realizability to make use of disjunctive properties on their behalf for the *reduction* of M* to P. Disjunctive properties even as cases of homeostatic property clusters are stable and recurrent enough to enable for reduction by means of *local identities*. Thus, Hunger as general phenomenon is realized differently by different species (Smith and Grueter, 2022). However, this does not hinder Hunger being locally realized across species such that we can identify even recurrent property clusters specific for *dog hunger*, *human hunger*, *mollusk hunger*. Hence, in this sense, disjunction of realization is compatible with reduction *by identification* between e.g., dog hunger, human hunger, mollusk hunger and their corresponding *local* HPCs. Furthermore, local reductionism is in a sense even compatible with non-reduction to overarching types by means multiple realization; dog hunger, human hunger and mollusk hunger are realized by a *multiplicity* of HPCs typical to their species. The multiple realization of a

higher-order property is just a *local phenomenon* carried out by recurrent yet differing sets of disjunctions being *locally realized*. In general, if multiple realizability allows for both reductionism and non-reductivism then conceptually speaking it is a contradictory thesis. As a corollary, if higher-level properties are not grounded in identities but in non-reducing disjunctions representing brute facts of nature, then the risk is imminently big that higher-level properties are causally speaking, epiphenomenal. In other words, if we want to secure the causal effectiveness of properties like M* we better do that by virtue of identifying the causal powers of M* with those held by P, even in those cases where P represents examples of HPCs which by virtue of their causal and constitutive relevance surpass the condition of mere brute facts. It is only to the extent that higher-level properties are identical with their disjunctive realizers at the local level of realization typical for a member X of a species S that higher-level properties typical for X are causally relevant at all. Two questions need to be answered to counter the critics. First, are homeostatic property clusters really cases of natural kinds and so on what bases? Second, are homeostatic property clusters then the sort of natural kinds that would be relevant for avoiding type identities even of the local sort as described above and can they consequently still grant causal efficacy to the phenomenon they realize?

5. Homeostatic property clusters as natural kinds

Two kind of conditions seems necessary to be fulfilled for clusters of properties to count as natural kinds; one is an epistemic and the other a metaphysical condition. Epistemically speaking, homeostatic property clusters must be possible to incorporate into explanatory models; they must contribute to our understanding of the world. Metaphysically speaking, homeostatic property clusters must reflect the structure of the world; in other words, they must have ontic weight. In order for homeostatic property clusters truly reflecting the structure of the world such that we can be sure they contribute to explanatory work in the sciences, they must first be identified *as* true natural kinds. Therefore, they must be stable and recurrent enough to be regarded as constituents of their phenomena both within and between individuals as in the case of biological structures; within and across species.

Furthermore, the clusters must be variable and flexible enough to allow for a variety of realizations of biological properties, avoiding the contradiction of allowing both reducibility and non-reducibility at the same time. Biological and psychological phenomena represent systemic properties whose nature and function are best reflected by structures the behavior of which echo the inherent plasticity and variability typical of biological systems, i.e., the condition of flexibility yet stability imposed on homeostatic property clusters (Bressler and Kelso, 2016). HPCs representing systemic properties such as those found in biology can accommodate to the synergy of both, on the one hand, reductively *local within-species* and on the other hand non-reductively *global* or *generally shared cross-species* properties, without contradiction. Thus, characteristic properties of a biological species such as morphology, reproductive isolation, predation, symbiosis and competition are locally defined by the species' local position within the available context of ecosystems to which it is adapted (Mazzocchi, 2008). At the same time the species properties transcend the limit of the ecological niche because they are *also* the result of *extended* patterns of adaptation in time through the process of evolution. Properties like robustness and generative integrity as articulated in the explanatory models of systems biology, trace actual conditions for the phenotypic expression of a particular behavior (Austin, 2016; Mason et al., 2015). Biological properties with both local and extended characteristics represent cases of homeostatic property clusters characterized by their ability to evolve new adaptations that require variability and plasticity at their level of realization (Duffau, 2006; Galván, 2010). Disjunction of properties would then be a possible grounding mechanism through which biological species develop new properties as an answer to the demands on developmental variability and evolutionary adaptation imposed by their environment. Further, homeostatic property clusters are by means of the variability, permeability and causal relevance characteristic of their structure and behavior cases of components implementing the characteristics of disjunctive property sets. Thus, if homeostatic property clusters basically correspond to the defining characteristics of sets of properties held together by disjunction and if we are willing to concede that disjunctive properties are cases of pluralistic natural kinds then at least by conceptual implication, HPCs should count as examples of natural kinds as well.

What we have established so far is the plausibility of both disjunctive properties and HPCs representing cases of natural kinds tracking the causal network structures and compositional aspects of the phenomena they refer to. Following this line of thought, the fact then that homeostatic property clusters can appear in explanatory predicates and projectible generalizations secures the status of disjunctive properties as natural kinds because disjunctive properties correspond to cases of HPCs and vice versa. More importantly, many systemic-oriented sciences, such as biology and psychology, seek to understand phenomena that are pluralistic. The explanations and generalizations developed in these sciences are consistent with properties that can be recognized as HPCs cases. Disjunctive properties as examples of HPCs can then appear as candidates of natural kinds and enter into realizing relations operating at the level of psychological and biological properties. The further implication is that property disjunction as a physical basis of realization would also block reduction by identities. This way we can answer the first horn of the dilemma: how to use disjunctive properties to provide irreducibility without sacrificing explanatory power. HPCs are locally stable clusters of realizing properties which in addition with their open-ended and context sensitive character still allows for the identification of causally relevant patterns of behavior globally extended over time. Thus, HPCs are even globally identifiable clusters of realizing properties whose instantiation reflects the historicity and iterability of evolutionary processes. From these patterns of behavior truly scientific explanations can be abstracted and developed as in the case of evolutionary and developmental biology (Gilbert, 2016; Watson et al., 2016). The same characteristics apply to disjunctive property sets when used in the recognition of realization relations. However, we must answer a further objection: granted that disjunctive properties have their own status as natural kinds as represented by the function and structure of HPCs, still, they cannot figure into *strict* lawful regularities and provide for the kind of theoretical generalizations that reflect those regularities. As the critics say, it is precisely the condition of "lawless" variability imposed on the definition of property disjunction as a realization relation that prevents its extension into lawful regularities. Consequently, disjunctive properties lack the metaphysical robustness to serve as the foundational basis for explanatory models in sciences like psychology.

Furthermore, this limitation results from the intrinsically *non-regular* nature of disjunction in the characterization of psychological phenomena. In other words, by saying that disjunctive properties have an inherently "non-lawful character", critics highlight that these properties do not follow consistent, predictable patterns that can be reliably used to predict and explain psychological phenomena. As a result, if sets of disjunctive properties represent the presence of HPCs as realization bases, psychological properties based on explanatory yet non-regular homeostatic property clusters would also be rendered explanatorily wild. Disjunctive properties and so even HPCs realize psychological properties but *do not* project nomically and therefore Psychology as a science would lack a firm metaphysical ground on which base its explanations in the context of non-reductive *physicalism.* Therefore, we can doubt the adequacy of the use of property disjunction, even as an instance of HPCs, as a coherent explanatory strategy capable of producing testable explanations. One preliminary answer could be: Psychology is not a strict science in the sense that it implements a copy of the methodological tools and conceptual schemas used in basic sciences like Physics. Psychological methodology is not *primarily* concerned with identifying the existence of bridging laws by which it could be reduced to a more basic science like neuro-physiology by the discovery of identity relations. Given the complexity of human behavior and cognition, psychology is not either *strictly* restricted to establishing in every case of analysis the conditions for ceteris paribus requirements to obtain: it's often difficult to account for all relevant variables involved in the gestation and function of psychological phenomena specially those concerned with the processes being characteristic to social cognition (Kelly et al., 2019). This leads to the conclusion that ceteris paribus conditions are more challenging in psychology than in some of the other sciences (de Jong, 2002). According to the critics, the realization of psychological properties at the level of their physical disjunctive sets would also render psychological properties causally inert. In a causally closed physical world all the causal effects are derived from the activity of the realizing physical bases. It could be granted there is no obstacle to the use of disjunctive base properties to causally realize psychological properties as long as their causal work is *determined by* the activity of those base properties. However, what the critics seems to oversee is that

when considered at the level of systemic properties, psychological phenomena remain causally effective on their own. Systemic properties transcend local causal limitations imposed by their physical brain bases because they involve even boundary breaking historical and context sensitive dimensions (Burnston, 2016a, 2016b, 2021). Thus, mental states carry the conditions of their past instantiations into their present constitution, and the same conditions are prospective for their future configurations in a way that makes them causally relevant. Besides, the brain itself at a systemic level is causally speaking boundary-breaking with regard to its physical structure because of its capability to form predictive loops embedded in the boundaries between its inner computations and its environmental context (Friston, 2010; Friston and Kiebel, 2009; McCaffrey, 2015). Psychological properties are also inherently contextual, embedded, and open-ended (Clark, 2017; Nave et al., 2020). This does not stop psychologists searching for *causal explanations* to construct explanatory models of human behavior such as under the controlled conditions of rigorous experimental set ups. However, to recognize that the subject matter of psychology does not conform to the *strict* methodological generalizations of fundamental sciences like physics is one thing. To deny that psychology can provide explanations appropriate to the properties that are typical at its own level of inquiry is something quite different (Fodor, 1980). Evidence-based psychology seeks explanations of behavior and mental content in terms of mechanisms and/or processes that conform to both intersubjective and intra-subjective regularities from which generalizations can be derived. The existence of these mechanisms and regularities is the theoretical bedrock on which an understanding of people´s behavior rests. Nonetheless, the kind of properties used in the discovery and explanation of those mechanisms are proper to the field; they are context-dependent and malleable. Psychological properties and mental properties in general reflect strong contextual and developmental aspects (Cacioppo et al., 2008; Sarter et al., 1996).

6. Disjunctive realization and special sciences

Psychological explanations generalize at their own level of analysis because the properties used are *permeable* yet *stable enough* to constitute

kinds in themselves from which explanatory generalizations can be drawn (Brick et al., 2022). For example, the content and function of psychological constructs such as our attitudes and attributional styles reflect the contextual character of our thoughts, the contingent features of our emotions, and the embedded conditions of our behavioral patterns. Psychological phenomena constitute *clusters of properties* held together by their systemic and synergetic causal effects. Psychological phenomena and their properties are *coarse* meaning they are broad in scope and multifactorial. Furthermore, the character of many psychological phenomena varies as a function of time, environmental and developmental effects. Thus, strictly speaking by their very *cluster character*, psychological properties cannot be reducible to neuroscientific type-to-type identities because these do not reflect the coarse character of psychological property variability. Psychological phenomena are realized by cluster properties and they enter into explanatory generalizations by virtue of their very disjunctive character. Thus, the coarse and contextual-developmental aspect of psychological kinds at their systemic level is matched at the realization level by the disjunction of coarse and pluripotential brain bases (Viola, 2021; Viola and Zanin, 2017). Explanations couched in the language of psychological theory link the broad character of the phenomenon to the coarse character of the factors and properties conforming to its realization. Psychology develops its own explanations for psychological properties without necessarily translating these explanations into the vocabulary of the sciences that typically deal with the physical instantiation of its properties. Thus, psychological explanations are coarse also because they involve many variables interacting with each other in often multidirectional causal ways. It is therefore that whenever matching psychological proprieties to their neurophysiological bases the coarse character of the psychological phenomenon must be considered. So, a right understanding of the coarse character of psychological phenomena can and should be reflected by the pluripotential properties of the instantiation bases; pluripotential brain states matching the multivariable character of psychological phenomena. Ontologically speaking then, psychological phenomena retain their irreducibility and explanatory autonomy without losing the power of being objects of theoretical generalization once we recognize that they represent homogenous and coarse properties. In other

words, psychological properties represent clusters of properties that, by virtue of being homeostatic, *also* possess the stability required to figure in recurrent causal explanations tracing ontic conditions. In other words, psychological phenomena and their properties as cases of homeostatic property clusters represent a realization by disjunction that still manages to preserve irreducibility and theoretical projectability for higher-order properties. Additionally, by regarding disjunctive bases as examples of HPCs, we establish their condition as natural kinds suitable for the explanatory purposes of special sciences like psychology and systems biology. Disjunctive bases as examples of HPCs then correspond to cases of *disjunctive realization* meaning that a variety of physical bases and their properties come together recurrently and differently to generate higher-order objects and their properties. These bases and their properties correspond to different patterns of realization that are stable enough to allow for explanatory generalizations as well. Moreover, the homeostatic character of the properties involved in disjunctive realization traces the existence of causally relevant patterns of behavior stable enough to indicate the presence of natural kinds; *disjunctive natural kinds*.

7. Systemic properties and their disjunctive realizations

Now, the critics might counter and say: granted that disjunctive realizations are explanatorily relevant reflecting the existence of disjunctive natural kinds in sciences like psychology. Still, this does not completely block reduction by identities because the pluralistic regularities observed at the level of psychological phenomena could be systematically translated to regularities at the level of their physical realizations, even in a pluralistic way. Thus, no matter how homeostatic and variable in character, properties and regularities described by special sciences such as psychology can still be explained in terms of the properties and regularities of realizing physical types. Reduction by identities is not prevented by disjunctive realization as conceived here. The question then is how much of the very *homeostatic character* of disjunctive kinds makes it possible to use that very stability and causal coherence to pursue actual identities. Thus, the critique goes that we could actually use HPCs as a backdoor strategy to provide the

necessary bridging laws that allow the reduction of psychological properties to neurophysiological ones (Bickle, 1992; Bickle et al., 2022). The ideal situation would be one in which the disjunctive properties of neuroscience *bridge the gap* between psychology and the brain, between the mental and the physical, giving rise to a unified picture of science including psychological science (Bickle, 1995). As an answer to this backdoor strategy, we need to consider the systemic nature of the brain and the nervous system as the biological bases responsible for the instantiation of psychological phenomena. Their systemic nature encompasses both synchronous and developmental as well as evolutionary processes that transcend the explanatory limitations imposed by purely physical, strictly law-based models of explanation. Thus, both synchronous and diachronic aspects of the brain's structure and functionality are represented by the systemic properties of the brain (Freeman, 2011). Synchronously, the brain is mechanistically constituted by the presence of intra-level non-causal components (Halina, 2017; Harbecke, 2015; Romero, 2015). Their very presence in the here and now in terms of their relations and activities constitutes the grounding bases of its structure and functionality (Darden, 2008; Machamer et al., 2000). Diachronously, these components are involved in the continuous making of casually relevant trajectories representing distal causes and effects. In other words, both structural and developmental aspects intersect to balance stability and variability in the development of the brain as a systemic organ. Because systemic structures such as the brain balance both stability and variability, they can be characterized by the coherent activity of homeostatic property clusters. Furthermore, such activity is indicative of causally informative relationships that are stable yet open enough to keep the system in a continuous flow of information and energy exchange with its environment. The backdoor strategy of locally founded reductionism is then blocked by the very systemic nature of the cluster kinds that condition explanations of psychological properties and their correlated brain states. Cluster types represent statistically based open-ended and time-dependent variables that defeat apparent property identities in terms of local species types (Lisman, 2017). The mere homeostatic character of disjunctive kinds cannot be used to trace identities in the way the critics mean. The backdoor strategy misunderstands the way "homeostatic character" is understood in sciences that

deal with both synchronic and diachronic aspects of realization, such as in psychology. Mental traits, for example, do not follow strict laws, but follow contextual and correlational patterns that vary from time to time within individuals, depending on the effects of phylogenetic, developmental, and environmental conditions affecting the subject. Such time dependent variability at the level of the mental is matched at the level of the realizing bases by means of variability and malleability among the constituents (Schulz and Hausmann, 2017). In other words, whatever the character of the realizing brain states and mechanisms responsible for the realization of higher-order cognitive properties, they must also be flexible enough to reflect the open-ended and dynamic nature of those properties (Kamaleddin, 2022). These conditions break the required symmetry between types standing in a reductive relation to each other. The crucial point for the establishment of type identity reductions among properties is that both types of properties instantiate more or less essentialist and monolithic types but the critiques implicitly misinterpret the pluralistic character of higher-order types and the pluripotential variability of their realizers. On the other hand, if they are willing to coincide that such pluralistic character must be matched by an equal plural character at the level of realizers then they are playing the game of the multiple realizability supporters and so entering into the realm on non-reductivism. Moreover, special sciences such as psychology have a different explanatory burden than the basic sciences because they must keep track of synchronous mechanistic properties as well as diachronous or historically extended properties. A psychological explanation of the effect of attitudes on people's behavior must not only consider conditions in the subject's proximal environment as well as distal factors but so even *reasons* conditioned by past events and experiences (Castelli and Tomelleri, 2008; Mezulis et al., 2004). Psychological agents act motivated by reasons and intentional content. Psychological properties own a historical and developmental dimension reflecting their permeability towards the environment; they are evolutionary and organismically based (Feinberg and Mallatt, 2016). The nature of the physical properties that realize them should at least functionally match this permeability, allowing for the necessary variability and diversity at the level of the constituents to play their realizing role. Disjunctive properties grounding psychological properties can then

instantiate both synchronous constitutive and diachronous developmental relations. Psychological properties are realized both by their brain-based constituents in a synchronous *here-and-now* manner and diachronously across spans of time by the shaping effects of *time extended* evolutionary mechanisms (Baciadonna et al., 2021). In other words, psychological properties are inherently dynamic. The constitutive relation between a higher-order phenomenon and the subvenient brain properties is one in which the latter aggregate temporally to instantiate the former. Explaining such patterns is best done by neuroscientific work based on the dynamics of brain states (Farmer, 2011). Thus, the neuroscientific bases realizing psychological properties will reflect the dynamical nature of those properties and by themselves show patterns of interchangeability, contextuality, and interdependent features of connectivity as well as functionality (Buzsaki, 2007; Nguyen et al., 2024). Dynamically interconnected brain networks showing patterns of synchronicities and distributed coupled oscillations seems to be the proper realizer candidates matching the temporal aspect of psychological properties (Deco et al., 2017; Demertzi et al., 2019). Explaining the evolutionary and developmental features of psychological kinds is currently done in the explanatory vocabulary of evolutionary and developmental psychology. Therefore, psychology as a science of the mind and behavior *depends* on the synchronous physical realization of neuroscientific kinds. At the same time, the diachronous character of psychological kinds reflects the open-ended character of those kinds whose behavior is best couched in the language of distributed networks representing dynamic activity patterns (Gallagher and Daly, 2018). Psychological explanations are representative cases of explanatory frameworks in the special sciences. Therefore, they do not necessarily follow the same methodology as the basic sciences, mainly because of the character of the properties they deal with. Such properties instantiate causal powers that are explained in terms of the dynamics of the systems that realize them. The role played by properties in the explanatory frameworks of sciences such as psychology and systems biology reveal the involvement of systemic and synergistic mechanisms and processes. Causally, such mechanisms and processes instantiate proximal and distal effects. These effects cannot be ordered into type-to-type identities that trace a one-to-one relationship. The establishing of such relationships is blocked by

contextual, evolutionary, and synergistic effects. Thus, "can we keep irreducibility by means of disjunction and still secure explanatory powers for special sciences?" As a primary answer, it seems plausible that we can have our cake and eat it too; we retain both irreducibility and theoretical projectability for the special sciences and their higher-order properties. Crucial to this effect is the characterization of the properties of the objects and phenomena treated in these sciences as genuine cases of natural kinds realized by equally genuine cases of instantiating natural kinds; disjunctive kinds. In this way we secure explanatory autonomy and value at the level of both systemic properties and their disjunctive realizations, complemented by irreducibility for higher-order properties. Still, critics may ask what sort of conditions apply to conglomerates of disjunctive properties to instantiate higher-level phenomena. By what means does disjunction of properties realize higher order properties? In other words, how does the *variability* implied by disjunction relate to the realization of irreducible properties?

8. Does variability or multiplicity require disjunction?

Thus, we need to answer the question of how variability or multiplicity relates to disjunction in order to establish irreducibility. We begin by observing that the general function of disjunctive properties is to realize higher order properties such as "desire for water" D(W). Further, we need to look more closely at what we mean by realization and the relation of disjunctive properties to functional performance. We have taken realization here to mean physical realization and, at first glance, to imply both functional occupant role implementation and material constitution. Realization is a variant of a metaphysical dependency relation offered as an alternative to identity theories (Polger and Shapiro: 2016 p.20). One understanding of realization then is to conceive of a set of physical properties P occupying a set of causal relations R such that the property D(W) is realized whenever the relevant set of physical properties in question, P, instantiates the set of causal relations R describing D(W). That is, D(W) is obtained at time *t* by virtue of P being in state R, at *t*. In other words, D(W) represents a causal role instantiation by P. What matters here is that my desire for water D(W) is functionally described as the relation that holds between P and R, and is

effectively realized whenever P executes R. What we have then is a view of realization that emphasizes the functional role description of higher order properties, where physical objects occupy functional roles; in this case of "desire for water": psycho-functionalism. More generally, a functional role description of D(W) highlights the set of inputs, internal states, and outputs representative of this higher-level property. Together, the members of this set describe the functional role of D(W) in relation to other similar properties in the broader network of psychological properties and their correlated behaviors. Thus, this set of inputs, internal states, and outputs captures a functional description **F** of D(W) in terms of the causal structure (C) of the behaviors and states characteristic of D(W). **F** is then the higher-order *abstraction* of the behavior represented by (C) and its corresponding states, hence F (my desire for water and my opening a soda can). This means that any suitably arranged set of objects and relevant properties similar to P instantiating causal structure (C) will suffice for the realization of D(W), because (C) is equivalent to P being in state R which in turn is the basis for the formation of **F**. In other words, it might be plausible that properly arranged pieces of cheese, i.e. suitably instantiating (C), will suffice for there to be D(W). This is obviously counterintuitive to common sense. The common-sense suspicion is based on the idea that cheese in general does not support mentality. Cheese has never been observed to instantiate a mental property; no hunk of cheese has ever screamed in pain when cut. There is something about the composition of cheese that excludes it as an appropriate realizer of mentality, however isomorphic its internal state arrangement might be to (C). Nevertheless, we should keep in mind the possibility that cheese, suitably arranged, can support other higher-order properties besides those traditionally associated with mentality, e.g., those properties inherent to its nature, such as its viscosity and hydrophobic structure. The objection to cheese as an appropriate realization of mental properties is based on the idea that something is *missing* from the microstructural-functional configuration of cheese the presence of which is necessary for giving rise to a subjective mental property. In general, this missing component would prevent cheese from exhibiting higher-order properties typical of other non-cheese systems and so there is no possibility for a chunk of cheese to realize the abstraction of a behavioral function such as **F**. This limitation refers to

a missing property or *the* missing property that if possessed by cheese, would allow cheese to instantiate D(W) or even cheese-pain. The same logic applies to disjunctive properties, since we could have disjunctive varieties of cheeses whose respective internal configurations would still lack the property by which D(W) is obtained. Alternatively, we override common sense intuitions and allow a liberal conception of realization as functional role occupancy even in the case of cheese. In other words, we reject the idea of a necessary property or set of properties, the presence of which would be crucial for the instantiation of properties like pain and desire. In other words, we allow that all objects whose internal configuration (C) represents F are sufficient for the realization of D(W) (Koskinen, 2020). If so, we are talking about conditions of *multiplicity* or variability for realizations to obtain, either by allowing a *restricted* set of key properties being distributed over the population of possible realizer systems, or by denying the existence of the missing key property and thus allowing for wildly different realizers in many different systems. The problem with the latter "wild" alternative is that multiple realization would then allow for an anything goes conception of realization, as long as it is variable yet stable enough to support realization. Variability in this sense seems to threaten the stability required by a higher-order property to maintain its defining characteristics across different domains. The background question is: wouldn't unrestricted variability in instantiating systems and their properties backfire on the coherence and homogeneity of the higher phenomenon being realized?

9. Ontological constitution, neurons and constrained disjunctive constitution

Alternatively, **P** and **R** represent sets of physical properties and their causal relations such that *not any* physical property nor *any causal relation* as proper part of those sets would be sufficient to realize **D(W)** and so obtain **F.** Only those properties and causal relations relevant in the sense of **P** instantiating a unique function ***G****, **F(*G**),** representing **D(W)** would do the job. Here **P** would stand for the relevant physical properties instantiating F**(*G**).** Those relevant properties represent then *a range* of limited constituents necessary for **D(W).** Thus, a second way of understanding

realization is as a case of ontological *constitution.* The property D(W) exists at all because it is physically constituted by realizers standing in a constitutive relation like that between a statue and the bronze of which it is made; the statue is its bronze and the bronze takes the form of the statue (Polger and Shapiro: 2016 p.20). What matters here is that it is physically constituted, or that it fulfills the condition of possessing a certain internal organizational structure relevant to its constitution (Melnyk, 2003). One important condition here is that a particular internal organization assumes the relevant functional role by virtue of its very constitutive nature. The statue being its bronze and the bronze being the statue cannot be anything other than the synchronous instantiation of the relevant functional roles occupied by the constituents themselves. The very presence of the constituents is the very presence of the functional roles that describe the phenomenon. The functionality automatically follows together with the arrangement of the constituent parts because the phenomenon is both structurally and functionally the typical arrangement of its constituent parts; it is the constituents that occupy the functional roles. Thus, in the framework of the mind-body relationship this version of physical realization is compatible with psycho-functionalism; the idea of *filler or occupant* functional role on the mind (Block, 2007). The mind certainly performs functions that can be captured by a functional description that traces the causal pathways of its relevant states. But these functions are also the result of its physical constitution being there in the first place; the mind is a function of constitution, not so much of causation; *the mind is in the flesh, not so much in the motion.* The unique function $\mathbf{F}(\boldsymbol{G^*})$, representing $\mathbf{D(W)}$ is there at all, as a matter of the right type of constituent relation being there at all. Functional role instantiation, then, is secondary but complementary to physical constitution, and it is the presence of the right kind of constituents that ensures that the right kind of function is performed at all. In other words, realization through constitution and realization through functional role occupancy are complementary sides of the same phenomenon; where there is constitution, there is function, and where there is function, there is constitution. However, this relationship is *ontologically asymmetrical* in the sense that the right type of constituent determines the right type of functionality, but not vice versa. The constitution, the bronze of which the statute is

made, is what matters ontologically. On the other hand, the functional description is what matters epistemically, because it informs about the special character of the relevant constituents and why there are constraints on the number of possible constituents. We could in principle have an approximate functional description of the phenomenon in question, allowing for the kind of variability of the constituents that physically realize the phenomenon. However, any kind of constituents would not suffice without misrepresenting the nature of the phenomenon itself. In filler role occupancy there are then certain *unique constituents* like neurons the role of which is to *embody* causal role (C) characteristic of D(W) for the realization of **F**. There is something about, and perhaps only about, neurons as constituents that makes them suitable for the realization of mental properties such as D(W). For all alternative objects other than neurons, such as silicon chips, which are in a one-to-one isomorphic relationship with the structural constitution of the biological neuron, something would still be missing for these objects to express D(W) (Shoemaker, 2009).

Thus, filler role occupancy describes the *epistemic consequences* of the *right type of constituents* being present to embody (C). It is by their very aggregate nature and instantiation as microconstituents that the right type of constituents such as neurons realize mental properties. What counts here is not so much the causal role (C) played by the neurons, but their spatial-temporal location which is what allows (C) to be embodied at all. In other words, what counts for the realization of D(W) by P is not so much that P performs R, but that P realizes D(W) through the very materialization of (C) by P; to realize is not primarily to cause, but to constitute. The condition of causal role fulfilment is secondary and subordinate to the condition of material constitution yet they are complementary. The functional role description *abstracts* from the constitutive level of the realizers the defining characteristics of the properties to be realized. The formation and constitution of D(W) by P is done by the type of mechanisms representing P synchronously and non-causally. The functional description in terms of causal specification represents the way the physical constituents "filler out" those mechanisms (Melnyk, 2018). Extending this idea to our understanding of the mind in a purely physicalist way, we can say that we hold beliefs about the world not because of the functions of the beliefs we hold, but because

of what our beliefs are made of; neurons synchronously constituting our beliefs right now in combination with their filling role properties. The right kind of material constitution then represents a sort of non-causal physical realization. There are then obvious selective constraints on the kinds of constituents that can realize higher-order properties, which in turn impose constraints on the kinds of *variability* that are open to the realization of properties like mental ones. Thus, from the point of view of multiple realization, constituents, as instances of physical realizations, impose limits on the kind of variability that can exist among realizers. Physical realization by constitution then represents a more conservative condition of realization because of its limitations on the scope of variability. What exactly are these limits? In other words, what counts as relevant but limited constitutive variability? The answer depends on the convergence of two conditions. One of them is that the evolutionary background conditions on Earth are unique. If the original conditions present at the beginning of time necessary for the chain of reactions leading to the realization of mental properties were to be repeated, there would be no exception to the rule that neurons and only neurons would support mentality by virtue of their distinctive biological properties. Moreover, if the original background conditions, including physical and biological conditions, were to deviate from the original conditions, then no mentality would ever have evolved because no neurons would ever have evolved either. Where there are no neurons, there are no functional role occupants of the right kind necessary for mentality to arise. The second condition implies the *property uniqueness* of the constituents e.g., the fact that neurons implement thinking is to be distinguished from the possibility of artificial systems based on other units than neurons implementing thinking. Neurons as such possess a property that makes them irrevocably unique as *units of mind* distinct from any other possible replica that could be designed; there is a special stuff, the constituent properties of evolved neurons, that makes them unique as realizers of mind. This is not to revive vitalism nor entelechies to explain the mind through realization. Rather, the unique combination of neurons as computational and constitutive elements of mind is an adaptive and autopoietic property, understood in terms of its evolutionary development and its ability to allow for and handle complexity in biological systems such as brains. Neurons, and only neurons as a type, are

the real stuff of the mind. This does not exclude that there may be *varieties* of neurons across species and within species, as in fact there are (Cauli et al., 1997). It is the richness of this diversity that leads to the diversity of cognitive abilities in living things, as long as this variability is limited by the constraints imposed by the evolutionary process (Güntürkün, 2012). The combination of these conditions (i.e., the evolutionary background conditions on Earth are unique and repeatable plus the property uniqueness of the constituents as seen in their ability to handle complexity through computational and constitutive properties) is also compatible with these conditions being exemplified by disjunctive properties as well, i.e., neurons possess and implement disjunctive qualities (Emery and Clayton, 2004). Neurons vary in types and within their types; there is variation at the individual scales of structural configuration as a result of the activities they enroll in, such as their synaptic activity and the arboreal networks they form and are shaped by (Clascá et al., 2012; Grafman, 2000; Lisman, 2017; Waschke et al., 2021). Neurons are plastic and implement learning at different scales of constitution (Toricelli et al., 2021). The conceptual association between neurons, homeostatic properties and realization by means of disjunction is the result of neurons having plasticity-based adaptative mechanisms and disjunctive properties characterized by both their structural and functional variability and malleability (Lisman, 2017). Nevertheless, it may still in the context of constituent realization be possible to have neurons representing cases of disjunctively-organized property clusters that realize species-specific local identities in a framework of conservative constituent realizability. Thus, *human pain* is in terms of constituents a local phenomenon the same way *mollusk pain* is constitutively local to the relevant mechanisms representing the relevant constituents; the functional networks specific to each species as implemented by their neuronal structures (Coninx, 2023) . The characterization of the constitutive condition is then a combination between metaphysical possibility regarding the variability of systems available for the instantiation of Pain, and empirical observations confirming both the existence of Pain as an across-species general property and the variability among its realizers. As such physical realization by constitution is conceptually compatible with multiple realization because constitution even in this conservative framework does not exclude variability when aspects of

plasticity and adaptability are taken into consideration. Let us call this condition for *constrained disjunctive constitution*. Thus, both metaphysical and empirical requirements are preserved for disjunctive constitution; metaphysically, one and the same property could be realized by a disjunction of relevant constituents, such as different types of neurons in aliens and other species. Empirically there is support for the kind of variability exemplified by disjunctive constitution being constraining but flexible enough to afford realization (Schouten and de Jong, 1999; Strappini et al., 2020). Still there is one important caveat to this story. Disjunctive properties instantiating functional roles as constituents in humans means sets of neurons standing to each other in *synchronous non-causal relations* to each other. Only their structural arrangement is metaphysically necessary for realization as long as they fulfill the requirement of being neurons. Disjunctive properties as instantiations of restricted constituents require then the presence of one and the same type of constituents for realizing higher-order properties such as D(W). Thus, because constrained disjunctive constitution represents *types* of properties, it is still open for those constitutive properties to express variability as *tokens* of that type. Anyhow, the restrictions imposed on realizers representing *the right type* of constituents makes this variant of disjunctive property realization explicitly conservative and chauvinistic. What physical realization by constitution apparently misses is the role played by the *dynamics* of the constituents; neurons do not realize mental properties only *as* neurons but essentially by the instantiation of their communication and connectivity patterns in brain networks (Demertzi et al., 2013; Fernandez-Espejo et al., 2012; Zhao et al., 2019). Thus, constitutive realization should be complemented by the dynamics of its components; their open-ended, synergistic, and distributed character exemplifying the degree of constraint and variability. The important question now is whether physical realization by constitution actually provides the right background for understanding the kind of variability relevant to multiple realization. Therefore, we are now in a position to evaluate the virtues of functional role occupancy and constitutive realization either in isolation or in combination. Several points we have already touched upon give us reason to doubt that physical realization by constitution alone can provide the kind of variability intended by multiply realizing systems. Following

multiple realization, this kind of variability must be sufficiently flexible yet stable enough to block reduction by type identities. First, in cases where multiple realizability is assumed, constitution is openly chauvinistic in the sense that it requires the existence of a unique property or a limited set of properties to be recurrent; this represents a *parochial attitude* towards realizability (Shoemaker, 2011). In line with multiple realizability, it seems implausible to believe that the diversity of the realizing bases by which nature may opt for the realization of higher-order properties would be so restricted.

Second, constitution explains higher-order properties like beliefs in terms of their material constituents or what seems to be equal to, in terms of their grounding mechanisms. This strategy misinterprets the very psychological character of our belief systems because its emphasis on constitution leaves out the *reasons, motives, purposes* and *contents* of our beliefs by means of which they are just what they are; beliefs. Third, a constitution view of realization regards variability as more or less static, *non-causally synchronous*, missing the role of the dynamics involved in the *formation and maintenance* of higher-order properties in psychology and biology. Fourth, realization by constitution is compatible with reduction through species-specific local identities. Fifth, constitution makes the functions played by the properties being realized dangerously epiphenomenal and relative to the causal efficacy of their material constitution. In other words, a closer look at the role of disjunctive constitution does not seem to serve the prospects of multiple realizability. Taken together, these points give us reasons to regard the kind of variability expressed by physical constituents as different in kind from the kind of variability intended by multiple realizability.

10. Amending constitution by making it dynamic

If disjunction of properties tracks variability it better does so by a different route than constitution *alone*. The suggestion is that a more reasonable alternative would be to *combine* the merits of constitution with the more ontologically relaxed claims of causal role functionalism. The emphasis here is on the *flow* of the connectivity pattern instantiating the functional

description, making the requirement on the constituents more relaxed without ignoring the existence of natural constraints. Thus, in theory, suitably arranged sets of constituents could implement the defining causal role (C) and do the same work, as long as they are sensitive to the variability in the flow of the processes necessary to create and maintain the phenomenon. In order to block the risk of inflation on the choice of constituents, at least some of the constraining characteristics imposed by constitution should be preserved and recognized as valuable: the variability of the causal connectivity patterns implemented by the realizing properties of the system states must be constrained by evolutionary and developmental background conditions and consequently by certain reasonable restrictions imposed on the material composition of the constituents. Therefore, many cases of realizers might do but not *any one* of them would do either; we do not obtain mental properties by combining varieties of cheese. Physical realization as constitution is about synchronously and non-causally grounding the existence of mental properties and nothing more. Once higher-order properties have so been grounded it is up to them to function in whatever way their special conditioning circumstances necessitate. They are merely grounded by their constituents, that´s all. Their functional profile is the result of embedded synergies. Further, physical realization by constitution does not mean higher-order properties necessarily inhere all their qualities from their realizers. Instead they develop own qualities and responses at their proper level of organization as a result of the adaptive needs specific to the system. Some structures and properties depending on their level of complexity would be strictly grounded in the common arrangement of the constituents giving rise to the higher-order property in question, others not (Gillett, 2010). Thus, the cutting index of diamonds is a higher-order property irreducible to its constituents; the carbon atoms in isolation by themselves do not cut and are not indexed by cutting values. It is their collective configuration that allows the diamond as a whole to possess the functional property of cutting through glass. Obviously, the presence of the carbon atoms and their atomic configurations *is* paramount for the diamond being a diamond at all, because no other elements than carbon can be arranged in the proper way to *be* a diamond. The explanation and description of the cutting index of diamonds is expressed by the suitable vocabulary of minerology.

Applying the same reasoning to the mind then, constitution as grounding relation would explain why mental properties exist at all but make no epistemological claims about why they are the way they are; making use of constitution as physical realization for mental properties is primarily about ontology not about epistemology. A further illustration would make the point hopefully clearer. In the statue-bronze example the statue and the bronze represent a mereological relation; the whole (the statue) is made by the parts (the bronze atoms) and the parts made up the whole. Still, the mereological relation misses the *idea* with the statue as such i.e., its representational content. The statue may represent Beethoven or a unicorn; the mereological relation is the same but the intended representation varies. The shape of the statue then is the *abstraction* out of which the constitution makes sense and may play a role in the network of contingent factors making up the networks of values and signifying properties it belongs to (Barbieri, 2011; Maran and Kleisner, 2010). In the same way, neurons constitute their mental properties both locally and collectively at the level of brain activity, but the idea or purpose with the constitution is to implement, among other things, thinking, and thinking is abstraction. Neurons and only neurons could be fundamental for the development of thinking minds. Nevertheless, their constituent character must express the presence of variability at the level of the thinking mind. The dynamics of the responsible constituents must counterbalance the dynamics of thinking as observed in its contextual and developmental aspects. From a scientific point of view, the grounding ontology of mental properties in the context of physicalism is best approached by the neurosciences but the proper explanation of mentality *as a function* of the mind implemented in behavior is best achieved by the science of psychology. Therefore, we can have a plurality of explanations for mental properties running in parallel encompassing different levels of representation without having reduction. Constituents *occupy* the functional roles of higher-order properties as the metaphysical conditions necessary for those roles to exist but nothing more besides that. Constituently speaking, any changes at the level of mental properties is reflective of changes in the grounding conditions provided by their brain constituents and their inherent dynamics.

Yet, the *dynamics* of those changes is characterized by the *abstracting behavior* of mental properties not by their brain constituents. It is in the domain of psychological science that the dynamics of these abstract models of behaviors are best approached and integrated into an explanatory framework that considers, among other things, the context-dependent and embodied nature of these properties (Albarracin and Pitliya, 2022; Buccino and Colage, 2022). The main virtue of constituents as grounding conditions is that their very constraining character provides the necessary stability conditions for non-reducible properties to obtain at all. Nothing changes this picture if we replace realizing bases by disjunctive properties acting as constituents, as long as these sets of disjunctions represent natural kinds in the way of HPCs intended here. In a sense, realization by constitution is compatible with a weak form of emergence, because constitution provides for higher-order properties that are derivable but whose proper functioning is not fully explainable by the constituent properties of their realizers alone. As a result, disjunction as constitution would be compatible with multiple realizability even under constrained conditions. The only caveat to keep in mind here is that constitution as disjunction is selective in the sense that only a limited range of properties allow for the kind of variability needed for multiple realization to follow. The nature of diamonds is limiting when it comes to disjunction, because no configuration other than those present in the atomic boundaries of carbon atoms present in diamonds would do the job. In the case of mental properties, we assume a more relaxed attitude towards their possible constituents and their variability. The running intuition is that mental properties are natural properties that are distributed across species, retaining unique properties in specific instantiations. The experience of hunger seems at least from a coarse point of view common to many creatures, and as such represents a unifying experiential as well as functional ground, but the specifics of the physiology that realizes hunger vary from species to species (Jourjine, 2017).

11. Conclusion

So far, we have established that if we assume that disjunction of properties represents natural kinds in terms of homeostatic property clusters,

then realization by disjunction finds a reasonable way out of the originally stated dilemma. We also see that disjunction as material constitution takes up restricted sets of realizing properties, but that any reduction by local identities can be blocked by these disjunctive properties grounding a form of weak emergence. We also observe that the explanatory value of irreducible properties lies in epistemological claims that are best addressed by relevant special sciences such as psychology, mineralogy, and biology. The abstractions from physical details are important as a way of understanding the phenomenon and the hallmark of special sciences is their capacity to produce abstractions with explanatory weight. Making use of constitutive disjunction does not preclude either the explanatory powers of higher-order sciences and their properties. Rather, the virtue behind constitutive disjunction lies in making it possible and even reasonable to expect the explanation of the functional aspects behind these properties couched in the vocabulary typical of those sciences in which *generalizable abstractions* of behavior play a fundamental role. Our analysis shows that: a) Higher-order properties can be realized by disjunction of lower-order properties without necessarily jeopardizing the requirement on *natural kindness* imposed on the latter. b) The kind of variability imposed by the multiple realizability thesis can be selectively restricted for certain properties such as mental ones being realized by neurons c) That the constraining effects of physical realization by constitution must also be counterbalanced by causal role instantiation in order to realize the *dynamical* characteristics of higher-order properties d) Explaining the nature of a higher-order property like the desire for water is in part achieved by the metaphysical grounding of its neuroscientific constituents but that the explanation of its functional role is best achieved by the explanatory powers of non-reducible sciences like psychology. e) Abstraction plays an important role in the realization relation.

In general, there seems to be a way out of the disjunctive dilemma that, according to the critics, affects multiple realizability: combining the virtues of property disjunction as cases of homeostatic property clusters with the selective constitution of realizers, complemented by functional role explanations. The variability appealed to in multiple realizability, which is necessary to ensure irreducibility, is not blocked by the use of disjunctive properties as the basis for this variability. Multiple realizability then stands as

an open alternative to the irreducibility of mental properties to their neuronal constituents; mental properties could be constitutively realized by only one kind of structures and their properties, i.e., neurons. Still, this picture is stationary and inherently "locationalist" enhancing the role of constituents as examples of mereological and non-casual synchronous properties. The dynamical aspect of the realization involved requires a liberalization of the role played by constitution in a complementary way that enhances global and distributed patterns of realization. Constitution must capture the dynamics of realization, and so this can be done both by having disjunctive constitution as one option and by having functional role implementation as its complement (Bressler and Kelso, 2016). The goal of this paper has been to work on the analytical flaws and virtues of multiple realizability. One of these vices, according to the critics, is the use of disjunctive properties to ground the variability of realization. The analysis has been intended to show that the plausibility of multiple realizability thesis still stands the test even when based on the use of disjunction as the subvenient ground for realization.

References

Albarracin, Mahault, and Riddhi J. Pitliya. 2022. "The Nature of Beliefs and Believing." Frontiers in Psychology 13: 1-5. https://doi.org/10.3389/fpsyg.2022.981925

Anderson, Michael L. 2016. "Neural Reuse in the Organization and Development of the Brain." Developmental Medicine and Child Neurology 58: 3–6. https://doi.org/10.1111/dmcn.13039.

Antony, Louise. 2003. "Who's Afraid of Disjunctive Properties?" Philosophical Issues 13 (1): 1–21. https://doi.org/10.1111/1533-6077.00001.

Austin, Christopher J. 2016. "The Ontology of Organisms: Mechanistic Modules or Patterned Processes?" Biology and Philosophy 31 (5): 639–62. https://doi.org/10.1007/s10539-016-9533-3.

Austin, Christopher J. 2020. "Organisms, Activity, and Being: On the Substance of Process Ontology." European Journal for Philosophy of Science 10 (2): 1–21. https://doi.org/10.1007/s13194-020-0278-0.

Baciadonna, Luigi, Francesca M. Cornero, Nathan J. Emery, and Nicola S. Clayton. 2021. "Convergent Evolution of Complex Cognition: Insights from the Field of Avian Cognition into the Study of Self-Awareness." Learning and Behavior 49 (1): 9–22. https://doi.org/10.3758/s13420-020-00434-5.

Balari, Sergio, and Guillermo Lorenzo. 2019. "Realization in Biology?" History and Philosophy of the Life Sciences 41 (1): 1–27. https://doi.org/10.1007/s40656-019-0243-4.

Barberousse, Anouk, Françoise Longy, Francesca Merlin, and Stéphanie Ruphy. 2020. "Natural Kinds: A New Synthesis." Theoria (Spain) 35 (3): 365–87. https://doi.org/10.1387/THEORIA.21066.

Barbieri, Marcello. 2011. "Origin and Evolution of the Brain." Biosemiotics 4 (3): 369–99. https://doi.org/10.1007/s12304-011-9125-1.

Bechtel, William, and Robert N. McCauley. 1999. "Heuristic Identity Theory (or Back to the Future): The Mind-Body Problem Against the Background of Research Strategies in Cognitive Neuroscience." In Proceedings of the Twenty First Annual Conference of the Cognitive Science Society: 67–72. https://doi.org/10.4324/9781410603494-17.

Bechtel, William, and Jennifer Mundale. 1999. "Multiple Realizability Revisited: Linking Cognitive and Neural States." Philosophy of Science 66 (2): 175–207. https://doi.org/10.1086/392683.

Bickle, John. 1992. "Multiple Realizability and Psychophysical Reduction." Behavior and Philosophy 20 (1): 47–58.

Bickle, John. 1995. "Psychoneural Reduction of the Genuinely Cognitive: Some Accomplished Facts." Philosophical Psychology 8 (3): 265–85. https://doi.org/10.1080/09515089508573158.

Bickle, John, André F. de Sousa, and Alcino J. Silva. 2022. "New Research Tools Suggest a 'Levels-Less' Image of the Behaving Organism and Dissolution of the Reduction vs. Anti-Reduction Dispute." Frontiers in Psychology 13: 1–12. https://doi.org/10.3389/fpsyg.2022.990316.

Block, Ned. 2007. Consciousness, Function, and Representation. The MIT press.

Boyd, Richard. 1991. "Realism, Anti-Foundationalism and the Enthusiasm for Natural Kinds." Philosophical Studies 61 (1–2): 127–48. https://doi.org/10.1007/BF00385837.

Boyd, Richard. 2021. "Rethinking Natural Kinds, Reference and Truth: Towards More Correspondence with Reality, Not Less." Synthese 198 (12): 2863-–903. https://doi.org/10.1007/s11229-019-02138-4.

Bressler, Steven L., and J. A. Scott Kelso. 2016. "Coordination Dynamics in Cognitive Neuroscience." Frontiers in Neuroscience 10: 1–7. https://doi.org/10.3389/fnins.2016.00397.

Brick, Cameron, B. Hood, Vebjorn Ekroll, and Lee de-Wit. 2022. "Illusory Essences: A Bias Holding Back Theorizing in Psychological Science." Perspectives on Psychological Science 17 (2): 491–506. https://doi.org/10.1177/1745691621991838.

Brigandt, Ingo. 2003. “Species Pluralism Does Not Imply Species Eliminativism.” Philosophy of Science 70: 1305–16. https://doi.org/10.1086/377409.

Bruggeman, Frank J., and Hans v. Westerhoff. 2007. “The Nature of Systems Biology.” Trends in Microbiology 15 (1): 45–50. https://doi.org/10.1016/j.tim.2006.11.003.

Buccino, Giovanni, and Ivan Colagè. 2022. “Grounding Abstract Concepts and Beliefs into Experience: The Embodied Perspective.” Frontiers in Psychology 13: 1–5. https://doi.org/10.3389/fpsyg.2022.943765.

Burnston, Daniel C. 2016a. “Computational Neuroscience and Localized Neural Function.” Synthese 193 (12): 3741–62. https://doi.org/10.1007/s11229-016-1099-8.

Burnston, Daniel C. 2016b. “A Contextualist Approach to Functional Localization in the Brain.” Biology and Philosophy 31 (4): 527–50. https://doi.org/10.1007/s10539-016-9526-2

Burnston, Daniel C. 2021. “Contents, Vehicles, and Complex Data Analysis in Neuroscience.” Synthese 199 (1–2): 1617–39. https://doi.org/10.1007/s11229-020-02831-9.

Buzsáki, György. 2007. “The Structure of Consciousness.” Nature: 267. https://doi.org/10.1038/446267a.

Cacioppo, John T., and Louis G. Tassinary. 1990. “Inferring Psychological Significance from Physiological Signals.” American Psychologist 45 (1): 16–28. https://doi.org/10.1037/0003-066X.45.1.16.

Cacioppo, John T., Gary G. Berntson, and Howard C. Nusbaum. 2008. “Neuroimaging as a New Tool in the Toolbox of Psychological Science: Commentary.” Current Directions in Psychological Science 17(2): 62–7. https://doi.org/10.1111/j.1467-8721.2008.00550.x.

Castelli, Luigi, and Silvia Tomelleri. 2008. “Contextual Effects on Prejudiced Attitudes: When the Presence of Others Leads to More Egalitarian Responses.” Journal of Experimental Social Psychology 44 (3): 679–86. https://doi.org/10.1016/j.jesp.2007.04.006.

Cauli, Bruno, Etienne Audinat, Bertrand Lambolez, Maria Cecilia Angulo, Nicole Ropert, Keisuke Tsuzuki, Shaul Hestrin, and Jean Rossier. 1997. “Molecular and Physiological Diversity of Cortical Nonpyramidal Cells.” Journal of Neuroscience 17 (10): 3894–906. https://doi.org/10.1523/jneurosci.17-10-03894.1997.

Chirimuuta, Mazviita. 2018. “Marr, Mayr, and MR: What Functionalism Should Now Be About.” Philosophical Psychology 31(3): 403–18. https://doi.org/10.1080/09515089.2017.1381679.

Clapp, Lenny. 2001. “Disjunctive Properties: Multiple Realizations.” The Journal of Philosophy 98 (3): 111–36. https://doi.org/10.2307/2678378.

Clark, Andy. 2017. "Busting Out: Predictive Brains, Embodied Minds, and the Puzzle of the Evidentiary Veil." Nous 51 (4): 727–53. https://doi.org/10.1111/nous.12140.

Clascá, Francisco, Pablo Rubio-Garrido, and Denis Jabaudon. 2012. "Unveiling the Diversity of Thalamocortical Neuron Subtypes." European Journal of Neuroscience 35 (10): 1524–32. https://doi.org/10.1111/j.1460-9568.2012.08033.x.

Colaço, David. 2022. "What Counts as a Memory? Definitions, Hypotheses, and 'Kinding in Progress.'" Philosophy of Science 89 (1): 89–106. https://doi.org/10.1017/psa.2021.14.

Coninx, Sabrina. 2023. "The Notorious Neurophilosophy of Pain: A Family Resemblance Approach to Idiosyncrasy and Generalizability." Mind and Language 38(1): 178–97. https://doi.org/10.1111/mila.12378.

Darden, Lindley. 2008. "Thinking Again about Biological Mechanisms." Philosophy of Science 75 (5): 958–69. https://doi.org/10.1086/594538.

Davidson, Donald. 2006. "Mental Events." In The Essential Davidson, 105–21. Oxford: Oxford University Press https://doi.org/10.1093/oso/9780199288854.001.0001.

de Jong, Huib. 2002. "Levels of Explanation in Biological Psychology." Philosophical Psychology 15 (4): 441–62. https://doi.org/10.1080/0951508021000042003.

Deco, Gustavo, Morten L. Kringelbach, Viktor K. Jirsa, and Petra Ritter. 2017. "The Dynamics of Resting Fluctuations in the Brain: Metastability and Its Dynamical Cortical Core." Scientific Reports 7 (1): 1–14. https://doi.org/10.1038/s41598-017-03073-5.

Demertzi, Athena, Andrea Soddu, and Steven Laureys. 2013. "Consciousness Supporting Networks." Current Opinion in Neurobiology 23 (2): 239–44. https://doi.org/10.1016/j.conb.2012.12.003.

Demertzi, Athena, Enzo Tagliazucchi, Stanislas Dehaene, Gustavo Deco, Pablo Barttfeld, Federico Raimondo, Charlotte Martial, et al. 2019. "Human Consciousness Is Supported by Dynamic Complex Patterns of Brain Signal Coordination." Science Advances 5 (2): 1–11. https://doi.org/10.1126/sciadv.aat7603.

Devitt, Michael. 2021. "Defending Intrinsic Biological Essentialism." Philosophy of Science 88 (1): 67-82. https://doi.org/10.1086/710029.

Dewhurst, Joe, and Alistair M.C. Isaac. 2023. "The Ups and Downs of Mechanism Realism: Functions, Levels, and Crosscutting Hierarchies." Erkenntnis 88 (3): 1035–57. https://doi.org/10.1007/s10670-021-00392-y.

Dosanjh, Ranpal. 2021. "Token-Distinctness and the Disjunctive Strategy." Erkenntnis 86 (3): 715–32. https://doi.org/10.1007/s10670-019-00127-0.

Duffau, Hugues. 2006. "Brain Plasticity: From Pathophysiological Mechanisms to Therapeutic Applications." Journal of Clinical Neuroscience 13 (9): 885–97. https://doi.org/10.1016/j.jocn.2005.11.045.

Edelman, Gerald. M., and Joseph A. Gally. 2001. "Degeneracy and Complexity in Biological Systems." Proceedings of the National Academy of Sciences of the United States of America 98 (24): 13763–18. https://doi.org/10.1073/pnas.231499798.

Emery, Nathan J., and Nicola S. Clayton. 2004. "The Mentality of Crows: Convergent Evolution of Intelligence in Corvids and Apes." Science 306 (5703): 1903–7. https://doi.org/10.1126/science.1098410.

Ereshefsky, Marc. 1992. "Eliminative Pluralism." Philosophy of Science 59 (4): 671–90. https://doi.org/10.1086/289701.

Esfeld, Michael. 2005. "The Causal Homogeneity of Biological Kinds." History and Philosophy of the Life Sciences 27 (3–4): 421–33.

Farmer, Simon. 2011. "Dynamic coordination in the brain: From neurons to mind." Journal of Psychophysiology, 25 (4): 201–2. https://doi.org/10.1027/0269-8803/a00005.

Feinberg, Todd E., and Jon Mallatt. 2016. "The Nature of Primary Consciousness. A New Synthesis." Consciousness and Cognition 43: 113–127. https://doi.org/10.1016/j.concog.2016.05.009.

Fernández-Espejo, Davinia, Andrea Soddu, Damian Cruse, Eva M. Palacios, Carme Junque, Audrey Vanhaudenhuyse, Eva Rivas, et al. 2012. "A Role for the Default Mode Network in the Bases of Disorders of Consciousness." Annals of Neurology 72 (3): 335–43. https://doi.org/10.1002/ana.23635.

Fodor, Jerry A. 1980. "Special Sciences (or: The Disunity of Science as a Working Hypothesis)." Synthese 28 (2): 120–33. https://doi.org/10.1007/BF00485230.

Francescotti, Robert M. 1997. "What Multiple Realizability Does Not Show." Journal of Mind and Behavior 18 (1): 13–27. http://www.jstor.org/stable/43853807.

Freeman, Walter J. 2011. "The Emergence of Mind and Emotion in the Evolution of Neocortex." Rivista Di Psichiatria 46 (5-6): 281–7. https://escholarship.org/uc/item/9w0672s3

Friston, Karl. 2010. "The Free-Energy Principle: A Unified Brain Theory?" Nature Reviews Neuroscience 11 (2): 127-38. https://doi.org/10.1038/nrn2787.

Friston, Karl, and Stefan Kiebel. 2009. "Predictive Coding under the Free-Energy Principle." Philosophical Transactions of the Royal Society B: Biological Sciences 364 (1521): 1211–21. https://doi.org/10.1098/rstb.2008.0300.

Gallagher, Shaun, and Anya Daly. 2018. "Dynamical Relations in the Self-Pattern." Frontiers in Psychology 9: 1–13. https://doi.org/10.3389/fpsyg.2018.00664.

Galván, Adriana. 2010. "Neural Plasticity of Development and Learning." Human Brain Mapping 31 (6): 879–90. https://doi.org/10.1002/hbm.21029.

Gilbert, Scott F. 2016. "Developmental Plasticity and Developmental Symbiosis: The Return of Eco-Devo." Current Topics in Developmental Biology 116: 415–33. https://doi.org/10.1016/bs.ctdb.2015.12.006.

Grafman, Jordan. 2000. "Conceptualizing Functional Neuroplasticity." Journal of Communication Disorders 33 (4): 345–56. https://doi.org/10.1016/S0021-9924(00)00030-7.

Güntürkün, Onur. 2012. "The Convergent Evolution of Neural Substrates for Cognition." Psychological Research 76 (2): 212–19. https://doi.org/10.1007/s00426-011-0377-9.

Halina, Marta. 2017. "Mechanistic Explanation and Its Limits." The Routledge Handbook of Mechanisms and Mechanical Philosophy: 213–24. https://doi.org/10.4324/9781315731544.

Harbecke, Jens. 2015. "Regularity Constitution and the Location of Mechanistic Levels." Foundations of Science 20 (3): 323–38. https://doi.org/10.1007/s10699-014-9371-1.

Hempel, Carl G. 1988. "On the Cognitive Status and the Rationale of Scientific Methodology." Poetics Today 9 (1): 5–27. https://doi.org/10.2307/1772885.

Jaworski, William. 2002. "Multiple-Realizability, Explanation and the Disjunctive Move." Philosophical Studies 108 (3): 298–308.

Jourjine, Nicholas. 2017. "Hunger and Thirst Interact to Regulate Ingestive Behavior in Flies and Mammals." BioEssays 39 (5): 1–11. https://doi.org/10.1002/bies.201600261.

Kamaleddin, Mohammad Amin. 2022. "Degeneracy in the Nervous System: From Neuronal Excitability to Neural Coding." BioEssays 44 (1): 1–14. https://doi.org/10.1002/bies.202100148.

Kelly, Michael P., Natasha M. Kriznik, Ann Louise Kinmonth, and Paul C. Fletcher. 2019. "The Brain, Self and Society: A Social-Neuroscience Model of Predictive Processing." Social Neuroscience 14 (3): 266–76. https://doi.org/10.1080/17470919.2018.1471003.

Khalidi, Muhammad Ali. 2018. "Natural Kinds as Nodes in Causal Networks." Synthese 195 (4): 1379–96. https://doi.org/10.1007/s11229-015-0841-y.

Koskinen, Rami. 2020. "Could We Really Be Made of Swiss Cheese? Xenobiology as an Engineering Epistemology for Biological Realization." ChemBioChem 21(18): 2591–4. https://doi.org/10.1002/cbic.202000063.

Lemeire, Olivier. 2021. "No Purely Epistemic Theory Can Account for the Naturalness of Kinds." Synthese 198 (12): 2907–25. https://doi.org/10.1007/s11229-018-1806-8.

Lisman, John. 2017. "Glutamatergic Synapses Are Structurally and Biochemically Complex Because of Multiple Plasticity Processes: Long-Term Potentiation, Long-Term Depression, Short-Term Potentiation and Scaling." Philosophical

Transactions of the Royal Society B: Biological Sciences 372 (1715): 1–11. https://doi.org/10.1098/rstb.2016.0260.

Machamer, Peter, Lindley Darden and Carl F. Craver. 2000. "Thinking about mechanisms." Philosophy of Science 67 (1): 1–25. https://doi.org/10.1086/392759.

Magnus, P. D, 2014. "NK≠HPC." The Philosophical Quarterly 64 (256): 471–7. http://www.jstor.org/stable/24672643.

Maran, Timo, and Karel Kleisner. 2010. "Towards an Evolutionary Biosemiotics: Semiotic Selection and Semiotic Co-Option." Biosemiotics 3 (2): 189–200. https://doi.org/10.1007/s12304-010-9087-8.

Mason, Paul. H., Juan. F. Domínguez Duque, Bodo Winter, and Andrea Grignolio. 2015. "Hidden in Plain View: Degeneracy in Complex Systems." BioSystems 128: 1–8. https://doi.org/10.1016/j.biosystems.2014.12.003.

Mazzocchi, Fulvio. 2008. "Complexity in Biology. Exceeding the Limits of Reductionism and Determinism Using Complexity Theory." EMBO Reports 9 (1): 10–4. https://doi.org/10.1038/sj.embor.7401147.

McCaffrey, Joseph B. 2015. "The Brain's Heterogeneous Functional Landscape." Philosophy of Science 82 (5): 1010–22. https://doi.org/10.1086/683436.

McCaffrey, Joseph B. 2023. "Evolving Concepts of Functional Localization." Philosophy Compass 18 (5): 1–17. https://doi.org/10.1111/phc3.12914.

Melnyk, Andrew. 2003. *A Physicalist Manifesto: Thoroughly Modern Materialism.* Cambridge University Press. https://doi.org/10.1017/cbo9780511498817

Melnyk, Andrew. 2018. "In Defense of a Realization Formulation of Physicalism." *Topoi* 37 (3): 483–93. https://doi.org/10.1007/s11245-016-9404-1.

Mezulis, Amy H., Lyn Y. Abramson, Janet S. Hyde, and Benjamin L. Hankin. 2004. "Is There a Universal Positivity Bias in Attributions? A Meta-Analytic Review of Individual, Developmental, and Cultural Differences in the Self-Serving Attributional Bias." Psychological Bulletin 130 (5): 711–47. https://doi.org/10.1037/0033-2909.130.5.711.

Minelli, Alessandro. 2019. "An Evo-Devo Perspective on Analogy in Biology." Philosophies 4 (1): 1-9. https://doi.org/10.3390/philosophies4010005.

Nave, Kathryn, George Deane, Mark Miller, and Andy Clark. 2020. "Wilding the Predictive Brain." Wiley Interdisciplinary Reviews: Cognitive Science 11 (6): 1–11. https://doi.org/10.1002/wcs.1542.

Nguyen, Tri D., Corey M. Magaldino, Jayci T. Landfair, Polemnia G. Amazeen, and Eric L. Amazeen. 2024. "From Cognitive Agents to Cognitive Systems: Theoretical, Methodological, and Empirical Developments of van Gelder's (1998) 'Dynamical Hypothesis." Topics in Cognitive Science 00 (2004): 1–19. https://doi.org/10.1111/tops.12725.

Noppeney, Uta, Karl J. Friston, and Cathy J. Price. 2004. "Degenerate Neuronal Systems Sustaining Cognitive Functions." Journal of Anatomy 205 (6): 433–42. https://doi.org/10.1111/j.0021-8782.2004.00343.x.

Noppeney, Uta, Will D. Penny, Cathy J. Price, Guillaume Flandin, and Karl J. Friston. 2006. "Identification of Degenerate Neuronal Systems Based on Inter-subject Variability." NeuroImage 30 (3): 885–90. https://doi.org/10.1016/j.neuroimage.2005.10.010.

Onishi, Yukinori, and Davide Serpico. 2022. "Homeostatic Property Cluster Theory without Homeostatic Mechanisms: Two Recent Attempts and Their Costs." Journal for General Philosophy of Science 53 (1): 61–82. https://doi.org/10.1007/s10838-020-09527-1.

Polger, Thomas W., and Lawrence A. Shapiro. 2016. The Multiple Realization Book. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198732891.001.0001.

Quian Quiroga, Rodrigo. 2020. "No Pattern Separation in the Human Hippocampus." Trends in Cognitive Sciences 24 (12): 994–1007. https://doi.org/10.1016/j.tics.2020.09.012.

Quian Quiroga, Rodrigo. 2023. "An Integrative View of Human Hippocampal Function: Differences with Other Species and Capacity Considerations." Hippocampus 33 (5): 616–34. https://doi.org/10.1002/hipo.23527.

Romero, Felipe. 2015. "Why There Isn't Inter-Level Causation in Mechanisms." Synthese 192 (11): 3731–55. https://doi.org/10.1007/s11229-015-0718-0.

Sarter, Martin, Gary G. Berntson, and John T. Cacioppo. 1996. "Brain Imaging and Cognitive Neuroscience: Toward Strong Inference in Attributing Function to Structure." American Psychologist 51 (1): 1–13. https://doi.org/10.1037/0003-066X.51.1.13.

Schneider, Susan. 2012. "Non-Reductive Physicalism Cannot Appeal to Token Identity." Philosophy and Phenomenological Research 85 (3): 719–28. https://doi.org/10.1111/j.1933-1592.2011.00534.x.

Schouten, Maurice K.D., and Huib Looren de Jong. 1999. "Reduction, Elimination, and Levels: The Case of the LTP-Learning Link." Philosophical Psychology 12 (3): 237–62. https://doi.org/10.1080/095150899105747.

Schulz, Jörg B., and Laura Hausmann. 2017. "The Malleable Brain – An Educational Review from Students to Students." Journal of Neurochemistry 142 (6): 788–9. https://doi.org/10.1111/jnc.14102.

Seitz, Rüdiger J., and Hans Ferdinand Angel. 2012. "Processes of Believing - A Review and Conceptual Account." Reviews in the Neurosciences 23 (3): 303–9. https://doi.org/10.1515/revneuro-2012-0034.

Shoemaker, Sydney. 2009. *Physical Realization*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199214396.001.0001.

Shoemaker, Sydney. 2011. "Realization, Powers, and Property Identity." Monist 94 (1): 3–18. https://doi.org/10.5840/monist20119411

Smart, J. J. C. 1959. "Sensations and Brain Processes." The Philosophical Review 68 (2): 141–56. https://doi.org/10.2307/2182164.

Smith, Nicholas K., and Brad A. Grueter. 2022. "Hunger-Driven Adaptive Prioritization of Behavior." FEBS Journal 289 (4): 922–36. https://doi.org/10.1111/febs.15791.

Soares, Deanne. P., and Meng Law. 2009. "Magnetic Resonance Spectroscopy of the Brain: Review of Metabolites and Clinical Applications." Clinical Radiology 64 (1): 12–21. https://doi.org/10.1016/j.crad.2008.07.002.

Sober, Elliott. 1999. "The Multiple Realizability Argument against Reductionism." Philosophy of Science 66 (4): 542–64. https://doi.org/10.1086/392754.

Strappini, Francesca, Marialuisa Martelli, Cesare Cozzo, and Enrico di Pace. 2020. "Empirical Evidence for Intraspecific Multiple Realization?" Frontiers in Psychology 11: 1–12. https://doi.org/10.3389/fpsyg.2020.01676.

Suthana, Nanthia, Arne D. Ekstrom, Michael A. Yassa, and Craig Stark. 2021. "Pattern Separation in the Human Hippocampus: Response to Quiroga." Trends in Cognitive Sciences 25 (6): 423–4. https://doi.org/10.1016/j.tics.2021.02.005.

Tonegawa, Susumu, Xu Liu, Steve Ramirez, and Roger Redondo. 2015. "Memory Engram Cells Have Come of Age." Neuron 87 (5): 918–31. https://doi.org/10.1016/j.neuron.2015.08.002.

Toricelli, Mariana, Arthur Pereira, Guilherme Souza Abrao, Helena Malerba, Julia Maia, Hudson Buck, and Tania Viel. 2021. "Mechanisms of Neuroplasticity and Brain Degeneration: Strategies for Protection during the Aging Process." Neural Regeneration Research 16 (1): 58–67. https://doi.org/10.4103/1673-5374.286952.

Viola, Marco. 2021. "Beyond the Platonic Brain: Facing the Challenge of Individual Differences in Function-Structure Mapping." Synthese 199 (1–2): 2129–55. https://doi.org/10.1007/s11229-020-02875-x.

Viola, Marco, and Elia Zanin. 2017. "The Standard Ontological Framework of Cognitive Neuroscience: Some Lessons from Broca's Area." Philosophical Psychology 30 (7): 945–69. https://doi.org/10.1080/09515089.2017.1322193.

Waschke, Leonhard, Niels A. Kloosterman, Jonas Obleser, and Douglas D. Garrett. 2021. "Behavior Needs Neural Variability." Neuron 109 (5): 751–66. https://doi.org/10.1016/j.neuron.2021.01.023.

Watson, Richard A., Rob Mills, C. L. Buckley, Kostas Kouvaris, Adam Jackson, Simon T. Powers, Chris Cox, et al. 2016. "Evolutionary Connectionism: Algorithmic Principles Underlying the Evolution of Biological Organisation in Evo-

Devo, Evo-Eco and Evolutionary Transitions." Evolutionary Biology 43 (4): 553–81. https://doi.org/10.1007/s11692-015-9358-z.

Wilson, Robert A, Matthew J Barker, and Ingo Brigandt. 2007. "When Traditional Essentialism Fails: Biological Natural Kinds." Philosophical Topics 35 (1/2): 189–215. http://www.jstor.org/stable/43154503.

Wilkins, John S. 2013. "Biological Essentialism and the Tidal Change of Natural Kinds." Science and Education 22 (2): 221-40. https://doi.org/10.1007/s11191-012-9450-z.

Zhao, Tong, Yiqian Zhu, Hailiang Tang, Rong Xie, Jianhong Zhu, and John H. Zhang. 2019. "Consciousness: New Concepts and Neural Networks." Frontiers in Cellular Neuroscience 13: 1–7. https://doi.org/10.3389/fncel.2019.00302.

Organon F 31 (3) 2024: 278–308
https://doi.org/10.31577/orgf.2024.31304

ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

Panpsychism: A Meta-View in the Philosophy of Mind

Artur Kosecki*

Received: 26 November 2023 / Revised: 19 July 2024 / Accepted: 22 July 2024

Abstract: In this article, my aim is to present panpsychism as a meta-view in the philosophy of mind rather than as a position that can be juxtaposed with leading positions such as materialism and dualism. I argue here that proponents of some versions of dualism, dual-aspect theory, some non-standard forms of physicalism, or idealism may be guided by the assumptions of panpsychism as a meta-view. For example, the literature includes positions such as Chalmers' naturalistic dualism, Strawson's physicalist panpsychism, and Sprigge's idealist panpsychism, along with Nagel's remarks on dual-aspect theory. I argue that panpsychism, as a meta-view, provides a framework within which to analyze how these positions address the mind-body problem. Consequently, I conclude that the solution to the mind-body problem itself remains neutral toward these positions. Instead of focusing on the elaboration of these metaphysical positions, attention should be directed toward the crucial issue for panpsychism: the combination problem.

Keywords: Combination problem; meta-philosophy; mind-body problem; panpsychism.

* University of Szczecin
https://orcid.org/0000-0002-3486-9404
University of Szczecin, Institute of Philosophy and Cognitive Science, ul. Krakowska 71-79, Szczecin, Zachodniopomorskie, Polska (Poland)
artur.kosecki@usz.edu.pl

1. Introduction

Panpsychism is typically understood as a metaphysical position in which mentality is considered fundamental and ubiquitous in the natural world (see Goff et al. 2022). However, if this description is taken literally, it might imply that everything in existence possesses a mind, consciousness, or experiences, including entities such as atoms or the Taj Mahal. Nevertheless, my objective is to present panpsychism as a meta-view rather than a position within the philosophy of mind, such as materialism[1] or dualism.[2] The

[1] In this context, it is important to distinguish between materialism and physicalism. Materialism has an older philosophical tradition than physicalism, as the premises of this stance were articulated well before the emergence of modern physics. The materialist position appeared in antiquity, with early philosophers like Democritus positing that the world is constructed from a single substance, namely matter. Contemporary physics, however, tends to speak not of matter *per se*, but of energy or mass—quantities that can be measured. Nevertheless, most philosophers today equate materialism with physicalism. In this article, I propose understanding materialism from a narrower viewpoint, asserting that everything that exists has a physical nature and that mental terms (e.g., "pain") reduce to physical terms (e.g., "C-neuron stimulation"). Physicalism, by contrast, is a broader position that allows the acknowledgment that not all that exists can be reduced to physical states or physical terms. This leads to distinctions such as reductive physicalism and non-reductive physicalism. For instance, a physicalist might accept mental states' supervenience on physical states, exemplifying non-reductionist physicalist stance. A good example of a non-reductionist physicalist position is Davidson's (1970) anomalous monism, which ontologically assumes the physical nature of the world but argues against the possibility of formulating psychophysical laws, thereby preventing the reduction of mental terms to physical terms.

[2] Skrbina (2017) seems to make a similar observation in his book *Panpsychism in the West*, when he writes about panpsychism as a meta-theory: "First, panpsychism is a unique kind of theory of mind. Its central feature is not that it examines or describes mind *per se*—although many panpsychists do this—but rather that it argues for a widespread or universal extent. In this sense it is a higher order theory, a meta-theory, of mind. It is a theory about theories. It simply holds that, however one conceives of mind, such mind applies to all things" (Skrbina, 2017, 3). Similarly, in the introduction to the collected work *Panpsychism: Contemporary Perspectives*, Brüntrup and Jaskolla contend that panpsychism has numerous variants and can be

literature discusses perspectives such as Chalmers' (1996) naturalistic dualism, Strawson's (2017a, 2020) physicalist panpsychism, and Sprigge's (1983) idealist panpsychism, along with Nagel's (1986) remarks on the dual-aspect theory. These positions illustrate that it is possible to adhere to the tenets of panpsychism while simultaneously embracing some versions of dualism, something I call a non-standard form of physicalism,[3] idealism, or dual-aspect theory.

These positions may therefore be based on what I would term a shared meta-metaphysical statement. Consequently, the question arises as to whether the potential solution to the mind-body problem is neutral with respect to these positions in the philosophy of mind. It is possible to represent dualism or a non-standard form of physicalism within panpsychism as a meta-view and have a similar understanding of the mind-body problem and its potential solution. Hence, panpsychism can be conceived as a meta-view that provides a unified framework for analyzing the mind-body problem across the aforementioned positions.

Here is the roadmap. In section 2, I will introduce the standard of materialism and dualism as it is typically understood in the philosophy of mind and demonstrate that supporters of these positions only partially address the mind-body problem. Therefore, some propose in the literature that panpsychism offers a position that can adequately respond to the mind-body problem.

In section 3, I will indicate that panpsychism should not be regarded as an alternative to the standards of materialism and dualism. Rather, it should be considered a meta-view that encompasses various positions in the

consistent with perspectives such as absolute idealism or substance dualism (Brüntrup and Jaskolla 2017a, 1).

[3] However, it should be noted that this is not in the standard understanding as typically presented in the literature, as I described in footnote 1. More specifically, it refers to Russell's (1927) assumptions outlined in *Analysis of Matter*, which are also discussed in section 4.3. Assuming that being a standard physicalist involves accepting supervenience, Davidson, although a non-reductionist physicalist, qualifies as a physicalist in the standard sense, not in the non-standard sense. Meanwhile, within the framework of non-standard physicalism that draws on Russell's assumptions, one posits the existence of mental properties as 'intrinsic' physical properties at the fundamental level.

philosophy of mind. I will note that the shared meta-metaphysical statement they adopt can be expressed through various ontologies.

I will also describe the combination problem that arises from adopting such a meta-view and point out that this problem results from what I term a bottom-up explanation.

In section 4, I will analyze the metaphysical positions that the bottom-up explanation adopts, illustrating their alignment with panpsychism as a meta-view. I will demonstrate how panpsychism provides a framework for the analysis of solutions to the mind-body problem within each of these positions.

In section 5, I will discuss panpsychism as a meta-view in the philosophy of mind, emphasizing that the solution to the mind-body problem is neutral of the positions outlined in section 4. Consequently, my conclusion is that instead of elaborating on the aforementioned metaphysical positions in an attempt to find a resolution to the mind-body problem, we should focus our attention on the combination problem as a crucial point in resolving the mind-body problem. I will also note that the proposal to solve the mind-body problem within the framework of panpsychism as a meta-view, and the combination problem, stems from the recognition that physical theories provide us with a limited description of reality. Section 5 will also address potential criticisms of my proposal for panpsychism as a meta-view.

2. The mind-body Problem: the standards of materialism and dualism[4]

In the literature on the mind-body problem, authors often highlight two primary concerns: (a) the problem of phenomenal consciousness[5] and (b)

[4] An elementary introduction to the positions in the metaphysics of mind was written by Levin (2022). In the article, I will outline only the standard materialism and the standard dualism in order to demonstrate the difficulties encountered in attempting to resolve the mind-body problem within these positions.

[5] In my article, when using the term "phenomenal consciousness," I refer to the concept made famous by Nagel (1974) in his essay "What is it Like to be a Bat?" I

the problem of causality (see Crane 2000, 169; Westphal 2016, 1–12). Regarding issue (a), an ongoing debate centers on the feasibility of reducing mental phenomena to physical states or reducing mental terms to physical terms, which also encompasses a broader attempt to reduce functional and intentional states to physical states.6 It turns out that addressing both (a) and (b) is problematic within the standards of materialism and dualism.

I understand standard materialism as the view that all facets of the mind, such as consciousness and thought processes, can be fully explained through physical terms. It asserts that mental states are nothing more than physical states. Historically speaking, materialism has been the dominant position since the 1950s, following the publication of two seminal papers: Feigl's (1958) "The 'Mental' and the 'Physical'" and Smart's (1959) "Sensations and Brain Processes".

However, since the 1970s, there has been a trend in the philosophy of mind opposing materialism, philosophers presenting thought experiments that challenge the possibility of reducing phenomenal states to physical states or terms (Chalmers, 1996; Jackson, 1986; Kripke, 1980; Nagel, 1974).[7] In other words, critics highlight the difficulties of capturing (a) the phenomenal character of consciousness within the standard materialism.

understand this type of consciousness as the subjective, first-person quality of experiences. According to Nagel, phenomenal consciousness entails that there is something it is like to be a specific organism; it encompasses a subjective character of experience that is inherently accessible only from the first-person perspective. However, in *The View from Nowhere*, Nagel points out that a similar issue was previously addressed by Sprigge and Farell (Nagel 1986, 15).

[6] In this article, I use "term" when referring to terms from psychological (e.g., "pain") or physical theories (e.g., "C-neuron stimulation"). However, I use "states" in an ontological context to refer to mental states (e.g., beliefs, desires, or sensations of pain) or physical states (e.g., neuronal activity in the brain).

[7] Significant contributions were made by Nagel's (1974) distinguished essay "What Is It Like to Be a Bat?" and Jackson's (1986) paper "What Mary Didn't Know." Kripke offered a significant critique of materialism in his *Naming and Necessity* (1980) by proposing his modal argument, and Chalmers' (1996) eminent book *The Conscious Mind* reignited the debate by renewing the modal argument in the form of the thought experiment of the Zombie.

On the other hand, those within materialism have doubts about the causal relation (b); they argue that its justification cannot be satisfactorily explained within the standard dualism. Generally speaking, the standard dualism posits that the mind and body are fundamentally distinct entities, meaning that mental states exist independently of physical states.[8] In this debate the principle of the causal closure of the physical world and the argument of causal overdetermination play crucial roles (Kim, 1993, 1998). According to this argument, for the occurrence of any physical states in the physical world, another physical state is sufficient; there is no need to appeal to a mental state as causally interacting with a physical state.[9]

Consequently, the standards of materialism and dualism only partially address the mind-body problem when understood in the context of (a) and (b). Thus, the choice between materialism and dualism leads to the following dilemma:

> **Dilemma (Dil.)**: We can either acknowledge the existence of phenomenal consciousness, which is not reducible to physical states or physical terms, as argued by standard dualism, or we can affirm that causal relations occur solely at the physical level, aligning with the standard materialist stance, while simultaneously reducing phenomenal consciousness to physical states.

It can therefore be seen that, on the basis of the core tenets of standards of materialism and dualism, an adequate resolution to the mind-body problem appears unachievable. The differences between materialism and dualism in their attempts to address the mind-body problem described above can be summarized as follows:

---

[8] When referring to the standard dualism, I primarily mean substance dualism, which posits the existence of two distinct entities. Substance dualism is difficult to reconcile with panpsychism. Panpsychism aims to integrate the mental and the physical. The literature also discusses property panpsychism, which, as demonstrated by the example of naturalistic dualism in section 4.1, fits within the framework of panpsychism as a meta-view outlined in section 3.

[9] The principle of the causal closure of the physical world served Kim (1993, 1998) in arguing against non-reductive physicalism, but it can also be used against the assumptions of dualism.

Standard materialism. Its proponents argue that phenomenal consciousness can be reduced to physical states or terms from physical theory.
***Advantage*:** An explanation of the causal relationship between mental state and physical state is obtained.
***Disadvantage*:** It reduces the role of phenomenal consciousness.[10]

Standard dualism. Its representatives argue that phenomenal consciousness is irreducible to physical states or terms from physics.
***Advantage*:** This position preserves phenomenal consciousness as something that actually exists and cannot be reduced to physical states or terms from physics theory.
***Disadvantage*:** There is a difficulty in adequately explaining the causal relations between mental states and physical states.

Consequently, in light of the difficulties that standards of materialism and dualism encounter in resolving the mind-body problem, panpsychism has recently attracted renewed academic interest.[11] The works of philosophers such as Goff (2017a, 2019) and Skrbina (2017), as well as edited volumes by Blamauer (2011a), Brüntrup and Jaskolla (2017b), Goff and Moran (2022), and Seager (2020), serve to illustrate this trend. In the next section of this article, I will describe the motivations from the literature that support panpsychism as an alternative to the standards of materialism and dualism.

[10] It is worth noting that some materialists would not consider this a disadvantage. For instance, Chalmers (2003) delineates distinctions within materialism. A type-A materialist contends that reductive explanations sufficiently account for all things that require explanation. On the other hand, a type-B materialist is prepared to accept an epistemic gap, as highlighted by anti-reductive arguments (Chalmers, 1996; Jackson, 1986; Kripke, 1980; Nagel, 1974), though this acknowledgment does not imply an ontological gap, since reality is fundamentally physical in nature. Consequently, the interpretation of the mind-body problem proposed in this article may be accepted only by certain materialists who believe there is a problem with the existence of phenomenal consciousness.

[11] Panpsychism appears to have a longer tradition than materialism and dualism, having been present in philosophy for 2,600 years. Skrbina (2017) wrote a brilliant monograph on the history of panpsychism and its contemporary challenges.

2.1. Panpsychism as an alternative for the standards of materialism and dualism

One philosopher who has considered the panpsychism position as a notable alternative to the standards of materialism and dualism in recent years is Chalmers, who wrote in his article "The Combination Problem for Panpsychism" (2017a, 179):[12]

> … that it promises to share advantages of both materialism and dualism and the disadvantages of neither. In particular, it can respect both the epistemological intuitions that motivate dualism and the causal intuitions that motivate physicalism.

The epistemological intuitions behind the aforementioned acceptance indicate that mental states are fundamental and irreducible. Thus, some proponents within the panpsychism position argue for the inclusion of phenomenal consciousness in the physical world (Chalmers, 2017b; Strawson, 2008b). From an ontological perspective, this forms the basis for the argument that, as mental states are considered fundamental and irreducible, they can participate in causal relations with physical states (Rosenberg 1996, 2004, 2017).

In the philosophical literature, the motivations for adopting panpsychism based on these grounds are termed "the explanatory argument for panpsychism" (Roelofs 2019, 14-15).[13] If proponents of the standard materialism cannot account for the complex phenomenology of living organisms solely in terms of matter, then it must be assumed that mental components already exist at the fundamental level of physical reality, either integrated with it or constituting its internal nature—otherwise we cannot

[12] In this passage, he refers to his article "Panpsychism and panprotopsychism" (2017b); in the article, he employs Hegelian argumentation for panpsychism, specifically pointing out that panpsychism is a synthesis of the thesis of materialism and the antithesis of dualism.

[13] Roelofs describes physicalism as the most common version of naturalism, which he understands as the position that there is one type of stuff that is regulated by fundamental laws. For him, among others, panpsychists would be anti-physicalist naturalists. He concludes that consciousness is a fundamental property of reality, similar to mass.

explain how conscious mind came to exist. Therefore, from an explanatory perspective, panpsychism provides a useful position.

Goff has also considered panpsychism to be an important alternative to both standards of materialism and dualism in discussions concerning the nature of the mind. In his article "The Case for Panpsychism" (2017b), he points out that:[14]

> Panpsychism offers the hope of an extremely elegant and unified picture of the world. In contrast to substance dualism (the view that the universe consists of two kinds of substance, matter and mind), panpsychism does not involve minds popping into existence as certain forms of complex life emerge, or else a soul descending from an immaterial realm at the moment of conception. Rather, it claims that human beings are nothing more than complex arrangements of components that are already present in basic matter. The only way in which panpsychism differs from physicalism is that the basic components of the material world also involve very basic forms of consciousness, from which the more complex conscious experience of humans and other animals derives (Goff 2017b).

In light of these considerations, some philosophers view panpsychism as an important alternative in debates about the mind, attempting to resolve problems arising from the assumptions inherent in the standards of materialism and dualism. From this perspective, human beings and other living organisms are seen as complex arrangements of components that contain mental elements at the fundamental level. This position obviates the need for both reductionism and theories of radical emergence in nature. Radical emergence proposes that emergent states are fundamentally unpredictable and irreducible to their components. Thus, mental states would have to emerge from a system that does not inherently contain mental elements.[15]

[14] The article is available on the website of the philosophical magazine *Philosophy Now*: https://philosophynow.org/issues/121/The_Case_For_Panpsychism

[15] In discussions on panpsychism, the argument against the occurrence of radical emergence within the metaphysics of mind debates is referred to as "the anti-emergence argument." I discuss this in section 4.3.

According to Goff, the assumptions described above suggest that panpsychism offers an elegant and unified picture of the world, combining the advantages of both materialism and dualism.

Thus, proponents of panpsychism offer an explanation for the existence of phenomenal consciousness, addressing problem (a), which the standard materialism struggles with in its own stance. Therefore, panpsychism does not endorse reductionism, as it recognizes genuinely existing mental states within its worldview. Panpsychists also clarify the occurrence of causal relationships between mental and physical states, thereby addressing issue (b), which standard dualism faces. Referring to Aristotle's principle of the golden mean, it can be asserted that it provides us with a middle path. Consequently, it can be argued that within panpsychism, a resolution to the mind-body problem (Dil.) may be achieved.

In the next section of the article, I will offer a definition of panpsychism as a meta-view in the philosophy of mind. Moreover, I will formulate a meta-metaphysical statement, which I maintain should be adopted by philosophers who guide different philosophical perspectives within panpsychism as a meta-view.

3. Panpsychism as a meta-view: terminological preliminaries

In this paper, I propose a definition of panpsychism as a meta-view in the philosophy of mind:

> **Definition**: Panpsychism is a meta-view within which it is posited that mental items exist at the fundamental level of reality, and due to their fundamental and irreducible nature, they enter into causal relationships with physical items.

In the aforementioned definition, the term "items" is used to emphasize that various ontologies can articulate the definition—namely, that at the fundamental level of reality, the components constituting reality may consist of properties, aspects, or events.

Thus, within panpsychism, there is allowance for an ontology of properties, which posits that fundamental reality is composed of elements that possess mental properties. Alternatively, an ontology of aspects may be

adopted, indicating that aspects constitute the fundamental level of reality. In the philosophy of mind, aspects are conceived as the manifestation of a given item, which can include both mental and physical aspects. In other words, if we have an elementary particle (e.g., an atom), it is accompanied by a mental counterpart by virtue of logical necessity. Therefore, there is a strong connection between the physical and the mental. Furthermore, an ontology of events can be considered as a model to describe the fundamental level of reality.

For this definition of panpsychism as a meta-view, it is important to emphasize the distinction between mental items at the fundamental level of reality and the developed mental life of humans or other living beings at a higher level of reality. Beliefs such as "Joe Biden is the President of the U.S.A." are not attributed to mental properties, aspects, or events constituted at the fundamental level. Nevertheless, such beliefs can be attributed to humans, as can phenomenal states (e.g., pain) experienced by both humans and other living organisms (e.g., dolphins).

However, for beliefs or phenomenal states to occur at the higher level of reality, there must be mental properties, aspects, or events at the fundamental level, as their appropriate composition produces the complex mental life of humans and other living organisms. In the context of panpsychism, the fundamental level of reality described above is also characterized as "smallism" (Coleman 2006). According to this view, entities are attributed specific states because they are composed of smaller items. For example, the nature and structure of a chair can be attributed to its composition of smaller components, such as atoms.

I argue that those who accept panpsychism as a meta-view ought to adopt the following meta-metaphysical statement in their metaphysical stance:

> **The meta-metaphysical statement (MS)**: The complex mental life of humans and other living organisms at a higher level of reality cannot fail to be composed of parts that lack mental items, and these elements constitute the fundamental level of reality.

In MS, "items" can be expressed using different ontologies within panpsychism as a meta-view. It can be said that the fundamental level of reality consists of mental properties, aspects, or events, which give rise to conscious minds only when appropriately constituted at higher levels of

reality. In other words, panpsychism as a meta-view in the philosophy of mind is neutral with regard to the ontology adopted.

In the following section, I will discuss the advantages and disadvantages within panpsychism as a meta-view as a proposition for solving the mind-body problem.

3.1. Panpsychism as a meta-view: the mind-body problem

In this article, I argue that panpsychism is not a stance that offers an alternative to materialism and dualism, but rather a meta-view that provides a framework for different positions in the philosophy of mind. Therefore, I believe that other positions also offer similar resolutions to the mind-body problem.

I will now present how, within the framework of panpsychism as a meta-view, the mind-body problem can be addressed.

> **Panpsychism as a meta-view.** This approach posits that the complex mental life of human beings and other living organisms can be explained in terms of a fundamental level where the components of reality contain mental items.
>
> ***Advantage*:** Phenomenal mental consciousness is assumed actually to exist; at the same time, an explanation is provided for the causal relations between mental and physical items.
>
> ***Disadvantage*:** How can distributed mental items at a fundamental level of reality produce the complex mental life of human beings and other living organisms at a higher level of reality?

Therefore, guided by D, which frames panpsychism as a meta-view in the philosophy of mind, we should provide the aforementioned response to the mind-body problem. Nevertheless, in addressing both problems (a) and (b), guided by D and MS, an issue arises, which I describe as a disadvantage—the combination problem (Seager 1995, 283).[16] Consequently, various positions in the philosophy of mind encounter the same problem. In the next section, I will outline this issue.

[16] Historically, James (1890) first articulated this issue in *The Principles of Psychology*. However, Seager (1995) referred to it as "the combination problem."

3.2. The combination problem

The combination problem is a challenge highlighted in the literature in which panpsychists must answer how mental items, which are simple, singular, and distributed at the fundamental level of reality, can combine to produce the complex mental lives of human beings and other living organisms at a higher level of reality (see Goff et al., 2022). For instance, Searle (2005), a prominent figure in the philosophy of mind known for his influential stance on biological naturalism, conveyed his apprehensions about panpsychism and the combination problem as follows:

> Aside from its inherent implausibility, pan-psychism has the additional demerit of being incoherent. I do not see any way that it can cope with the problem of the unity of consciousness. Consciousness is not spread out like jam on a piece of bread, but rather, it comes in discrete units. If the thermostat is conscious, how about the parts of the thermostat? Is there a separate consciousness to each screw? Each molecule? If so, how does their consciousness relate to the consciousness of the whole thermostat? (Searle 2005, 150).

It is worth noting that the growing interest in panpsychism has sparked discussions about the combination problem and its potential solutions (e.g., Coleman 2006, 2012, 2014, 2017; Goff 2009a, 2009b, 2017c; Roelofs 2019, 2020; Rosenberg 2004; Seager 2010, 2017). Nevertheless, in the article I will refer to Chalmers' paper "The Combination Problem for Panpsychism," where he presents a framework for contemporary debates regarding the combination problem. He distinguishes three forms of this problem in discussions about the mind (Chalmers 2017a, 182–185):[17]

(i) The subject combination problem.
(ii) The quality combination problem.
(iii) The structure combination problem.

[17] Other distinctions of the problem of combination found in the literature include Coleman's "internal" and "bridging" categories (Coleman 2017), Goff's "from above" and "from below" distinctions (Goff 2017c), and Roelofs' "hard" and "easy" problems (Roelofs 2020).

According to Chalmers, a comprehensive understanding of how mental life is produced requires addressing all variations of the combination problem (Chalmers 2017a, 184).

Problem (i) pertains to explaining how scattered micro-subjects combine to form a single macro-subject. This challenge can be illustrated using the metaphor of Lego bricks. The appropriate assembly of Lego bricks, each representing simple elements, results in a cohesive structure, such as a Lego brick-built building. Roelofs (2020) has contended that the "subject" problem of combination is the "hard" problem in panpsychism, while other variations represent "easy" problems that may be resolved through "phenomenological analyses," focusing the subject's attention on its own internal experiences (Roelofs 2020, 246).

Coleman (2014) explains the difficulties associated with (i) by indicating that subjects cannot combine because individual perspectives exclude one another. Specifically, it would lead to contradictions if one subject believes A while another believes not-A, and when they combine into a new subject, they would have to believe both A and not-A or the subjectivity of one of them would not contribute to the new subject. This presents a dilemma that needs to be resolved within (i).

On the other hand, Coleman (2012) also points out that the combination problem results from an assumption regarding the existence of subjects of experience and the conception of unity in terms of aggregates and constituent parts. He employs the "model metaphor" of the palette to describe this problem: just as colors are mixed on a painter's palette, a complex phenomenology of experience is formed from individual impressions (Coleman 2012, 157). He addresses the problem of combination in form (ii), where it is necessary to explain how a complex palette of phenomenal experiences (macro-qualities) arises from simple qualities (micro-qualities).

Issue (iii) pertains to Russell's (1927) stance on the philosophy of mind as presented in its source form, neutral monism.[18] If physical and phenomenal properties are deemed to represent the same reality, which we comprehend precisely through these properties, then it is presupposed that their structures are isomorphic. Yet, the macro-phenomenal structure is richer

[18] It is worth noting that this argument against Russell's monism was articulated by Stoljar (2001) in his article "Two Conceptions of the Physical."

and more phenomenologically complex than the micro-structure. Given this, macro- and micro-structures should be viewed as incommensurable. Such a perspective leans more toward a version of dualism, since, from this standpoint, we could argue that we are discussing two distinct structures that are not reducible to one another.

Thus, in addressing the combination problem, one must provide a complex response, such as detailing how (i) micro-subjects constitute macro-subjects. They must also elucidate (ii) how individual experiences combine to give rise to the complex phenomenology of the human organism and other animals. Additionally, they are required to address (iii), how a rich macro-phenomenal structure arises from a micro-phenomenal structure.

Therefore, solving both problems (a) and (b), and consequently Dil., within panpsychism as a meta-view, is not straightforward. In the next part of the article, I will describe how the combination problem stems from what I term a bottom-up explanation, which is accepted within different positions in the philosophy of mind. I argue that these positions are in alignment with the panpsychist perspective outlined in section 3. Thus, I will show how these positions fit within the framework provided by panpsychism as a meta-view. I will also indicate that, based on these stances, one must confront the combination problem.

4. Bottom-up explanation and panpsychism as a meta-view in the philosophy of mind

The combination problem arises from adopting what I have referred to as a bottom-up explanation. To express the intuitions behind acceptance of the meta-metaphysical statement presented in MS—namely, to explain the production of complex mental life in human beings and other living organisms—one must accept that the fundamental level of reality also contains mental items. In the literature, four models of bottom-up explanation can be identified in the philosophy of mind. Each model is based on a different ontology. These models attempt to explain how complex mental life is produced at a higher level from some mental items at the fundamental level of reality:

(M_1) In addition to the existence of fundamental physical properties, fundamental mental properties exist.

(M_2) Every fundamental constituent of reality has both a mental and a physical aspect.

(M_3) Every fundamental constituent of reality has both mental and physical properties.

(M_4) Mental events are fundamental, and there are no other beings at this level of reality.

Therefore, various ontologies, such as properties, aspects, or events, can be adopted to express the MS. Consequently, different bottom-up explanatory models (M_1–M_4) are assumed within various positions in the philosophy of mind.

4.1. Chalmers' naturalistic dualism

In the context of M_1, property dualism can be adopted, for example, which posits two types of property: mental and physical. Chalmers' naturalistic dualism is a contemporary example of property dualism. He delineates this position in his seminal work, *The Conscious Mind* (1996). While he acknowledges the principle of causal closure of the physical world, he argues, based on non-reductive reasoning such as his version of the conceivability argument, that there is not an *a priori* relationship between physical and phenomenal properties. This implies that laws grounded in physical theories fall short in fully explaining the conscious mind. Thus, to formulate a comprehensive theory of consciousness, Chalmers suggests the existence of new fundamental properties and laws separate from the laws of physics. He posits that mental properties should be viewed as fundamental properties similar to physical properties like mass or force.

Chalmers brands his viewpoint naturalistic, aligning it with the principle of causal closure of the physical world. Within this position, he suggests that phenomenal states might hold an epiphenomenal status. Importantly, to address the challenges associated with epiphenomenalism, Chalmers turns to the panpsychist perspective. He references Rosenberg's (1996) proposition, emphasizing that the mere presence of experience might enable causal relationships, and hints that addressing epiphenomenalism could introduce the counterintuitive concept of panpsychism (Chalmers 1996, 152).

However, Chalmers (1996, 297–299) also addresses the issue of panpsychism in his book in the context of his double-aspect principle of information. Within this approach, he examines "information" in relation to Shannon's concept (1948) as something that possesses both a physical and a phenomenal aspect. This refers to the double-aspect principle, which may be a consequence of the principle of structural coherence he adopts, dealing with the correspondence between subjective phenomenal states and states of consciousness (e.g., attention).[19] In this manner, he analyses the issue of panpsychism as an answer to the hard problem of consciousness:

> The view that there is experience whenever there is causal interaction is counterintuitive. But it is a view that can grow surprisingly satisfying with reflection, making consciousness better integrated into the natural order. If the view is correct, consciousness does not come in sudden jagged spikes, with isolated complex systems arbitrarily producing rich conscious experiences. Rather, it is a more uniform property of the universe, with very simple systems having very simple phenomenology, and complex systems having complex phenomenology. This makes consciousness less "special" in some ways, and so more reasonable (Chalmers 1996, 298).

In the above quotation, the so-called continuity argument is discussed, which posits that if lower forms of life possess certain, albeit most elementary, phenomenal experiences, it can be expected that consciousness as such is present, albeit to varying degrees, in all forms of life. This leads to a position suggesting that mental items might be fundamental and ubiquitous.[20]

Thus, in his philosophy, Chalmers examines M_1, property dualism (naturalistic dualism), or M_2, the dual-aspect theory (the double-aspect principle and the principle of structural coherence), within the broader context of panpsychism. Under these circumstances, panpsychism could serve as a meta-view from which Chalmers seems to explore the adoption of the

[19] Chalmers (1995) also discusses these principles in his seminal article "Facing Up to the Problem of Consciousness." However, this article does not address the issue of panpsychism.

[20] Therefore, some argue that if evolution is to proceed smoothly, consciousness must have been present in some form from the very beginning (e.g., James, 1890).

MS and refers to it as a possible approach to addressing the mind-body problems (a) and (b), thereby offering a potential resolution for the dilemma (Dil.). However, as mentioned, by adhering to these assumptions, we encounter the combination problem.[21]

4.2. Nagel's remarks on the dual-aspect theory

While Chalmers is often credited with popularizing panpsychism in debates on the mind during the 1990s, that it is worth noting the arguments for panpsychism currently discussed in contemporary metaphysics of mind were presented by Nagel in his 1979 essay "Panpsychism".

In this essay, Nagel (1979) argued that panpsychism could be one option for addressing the mind-body problem. It is important to note that the paper has already outlined the combination problem, which he understood as (i) the subject combination problem. He describes the difficulty of conceiving how mental states in complex organisms could arise from the protomental properties of dead matter. He points out this problem as follows:

> Yet they would have to be recompilable to form different points of view, for not only can a single organism have different experiences, but its matter can be recombined to form other organisms with totally different forms of experience (Nagel 1979, 94).

Hence, it seems that he implicitly postulates the MS within the panpsychism he delineates.

In his influential book *The View from Nowhere* (1986), Nagel discusses panpsychism, emphasizing that it emerges from adopting the double-aspect theory (Nagel 1986, 49).[22] However, it seems more accurate to say that Nagel adopts the framework of panpsychism to engage with the dual-aspect theory more profoundly. This theory strives to integrate both subjective

[21] The article on how property dualism leads to panpsychism was written by Blamauer (2011b). In the article, he references Chalmers but does not discuss his views in the context of the combination problem as it had not been developed at that time.

[22] Nagel wrote about panpsychism as a consequence of the dual-aspect theory, which he describes as having the "odor of something put together in the metaphysical laboratory" (Nagel 1986, 49).

and objective issues, encompassing first-person and third-person perspectives, as well as mental and physical aspects. This position presupposes M_2, and consequently MS. Therefore, such a theory should also address the combination problem.

What is puzzling, however, is that in his works, Nagel sometimes writes about proto-mental properties and at other times about panpsychism (Nagel, 1979, 1986, 2012), as if he does not distinguish between these two positions in his philosophy.[23] Another criticism of Nagel's philosophy of mind is that his approach provides an answer to problem (a) but lacks a properly formulated theory of mental causation (b) in relation to the mind-body problem.

4.3. Strawson's physicalist panpsychism

Regarding M_3, it is worth noting two important works by Strawson: "Real Materialism" (2008a [2003]) and "Realistic Monism" (2008b). In these papers, he emphasizes the realism of the mental within the physicalism. According to Strawson, to be a realist materialist, one must recognize that experiential being constitutes the intrinsic nature of physical beings. In the literature, this argument is known as the intrinsic nature argument for panpsychism. Furthermore, a significant motivation for adopting the panpsychist perspective is its critique of theories based on radical emergence. This argument is called the anti-emergence argument. One would present this argument in relate to the paper "Realistic monism" in the form of a question: how can one explain the experiential from something that is non-experiential?[24] Strawson concludes this way: "the existence of every real

[23] Thus, panprotopsychism is a position that states that at the foundational level, proto-mental properties exist, but these properties are potential, and only through their specific organization does a conscious mind emerge at higher levels. It is also worth noting that panprotopsychism can be endorsed by both neutral monists (Chalmers 1996, 2003) and proponents of physicalism (Strawson 2008b).

[24] However, it is important to emphasize that in the context of M_3, Strawson (2008a [2003]) suggests replacing "mental" and "physical" with "experiential" and "non-experiential." According to him, this division between mental and physical leads to a debate between materialism and dualism.

concrete thing involves experiential being even if it also involves non-experiential being" (Strawson, 2008b, 57).

In "Real Materialism," Strawson refers to Russell's (1927) view as it is presented in the book *Analysis of Matter*, where Russell emphasizes the use of mathematical apparatus in physics. According to Russell, this leads to the conclusion that physics does not provide a complete description of reality. When referring to physical theories, we are merely describing the mathematical dimensions of physical being and focusing solely on the structure of physical reality. Consequently, within physical theories, we do not attribute any characteristics to the intrinsic nature of physical being. However, when we assume that certain physical states are identical to mental states, we gain direct familiarity with their intrinsic nature through our own experiences.

Thus, our understanding of the physical may appear entirely different from what it truly is. Therefore, Strawson's redefinition of the term 'materialism' in his work "Realistic materialist" has the potential to lead to the consideration of panpsychism or panexperientialism[25] as plausible options in discussions about the mind (Strawson, 2008b, 71).

Hence, it can be inferred that Strawson's perspectives are in line with the M_3 model. The fact that a given component of reality is made up of experiential and non-experiential properties can be put in terms of the division into intrinsic and extrinsic properties, or non-structural and structural properties. The experiential would be an intrinsic or non-structural property, whereas the non-experiential would be an extrinsic or structural property.

It can therefore be concluded that, on the basis of his philosophy, Strawson adopts the meta-metaphysical statement MS while trying to resolve Dil. It can be posited that experiential properties (a), which may be considered as intrinsic or non-structural properties, would play a causal role (b). One can argue, within Strawson's position, that intrinsic properties enter into causal relationships with extrinsic properties. Rejecting the possibility of a

[25] The stance of panexperientialism states that everything that exists experiences or has the ability to experience (Skrbina 2017, 16). It is worth noting that Chalmers and Strawson treat panexperientialism as synonymous with panpsychism (Chalmers 2017b, 19; Strawson, 2015, 201). Panexperientialism was also adopted and defended by Griffin (1997, 1998). See section 4.4.

causal role for experiential properties would be tantamount to denying their irreducible and fundamental nature.

However, when confronted with the combination problem, Strawson regards it as a trivial problem that does not present any significant challenges (Strawson, 2017b, 100). What is more, he even admits that he does not know what a potential solution to this problem would look like.

Nevertheless, I have outlined the development of Strawson's position, which he recently termed "physicalist panpsychism" (2017a, 2020). It seems that panpsychism as a meta-view provides a framework within which what I refer to as non-standard physicalism—termed "physicalist panpsychism" by Strawson—integrates experiential and non-experiential properties.

4.4. Sprigge's idealist panpsychism

In this context (M_4), I will discuss the idealism of Sprigge, which can be understood as an attempt to integrate panpsychism with absolute idealism (McHenry 2010). Sprigge argued for a version of absolute idealism based on the panpsychist perspective, according to which reality consists of bits of experience combined into a coherent whole. He explored the framework of panpsychism as suitable for tackling the mind-body problem and delineating the nature of noumena in his book *The Vindication of Absolute Idealism* (1983).

When explaining Sprigge's views, it is worthwhile incorporating the differentiation between appearance and reality, which was introduced by the prominent British idealist Bradley (1893 [1969]). This differentiation elucidates that the physical world, described structurally, is only the world of appearance, while the true reality exists behind it—that is, experience (see McHenry 2010).

According to Sprigge, experience is directly known by us. Strawson puts forth a similar argument regarding the intrinsic nature of physical beings, as previously mentioned.[26] Additionally, it should be highlighted that Sprigge's philosophical beliefs align closely with process philosophy.[27]

[26] It is noteworthy that Strawson cites Sprigge as a proponent of physical pure panpsychism (Strawson 2020, 319).

[27] It should be noted that Strawson's philosophy of mind also leans toward a process philosophy (Strawson 2017b).

In his philosophy, Sprigge also delves into the ontology of events, which is typical among representatives of process philosophy. At a fundamental level, events might constitute an experience and form a larger whole. Whitehead[28] is widely regarded as the foremost representative of process philosophy and as a leading panpsychist in twentieth-century philosophy,[29] but his pupil, Hartshorne (1950), brought this position into the mainstream (Skrbina 2017, 228),[30] representing the perspective of "psychicalism" (Hartshorne, 1977).[31] However, Griffin (1997, 1998), in reference to the philosophers mentioned above, posited that experience exists at the fundamental level of reality; he believed that the most suitable term for this is "panexperientialism".[32]

Within Sprigge's idealism, it is worth noting that the metaphysical system aimed at synthesizing panpsychism and absolute idealism falls under category M_4. One might argue that this system is based on a meta-metaphysical statement MS. Furthermore, one can expect that based on idealistic panpsychism, we will address Dil. And, consequently, address (a) and (b). However, it also grapples with the combination problem, which may

[28] It is pertinent to highlight that Whitehead authored *Science and the Modern World* (1925), wherein he advocated a metaphysical perspective distinct from the tenets of reductive materialism and dualism.

[29] I t is worth noting that Whitehead himself did not use the term "panpsychism" in his philosophy. Cobb (2008) writes about this, introducing the term "panpsychism" in the entry "occasion of experience."

[30] It is worth mentioning that he wrote an encyclopedia article regarding "panpsychism" in the published work *Philosophical Systems* (Hartshorne 1950, 442-453).

[31] Also, in 1977, Hartshorne presented an argument for panpsychism in a significant article titled "Physics and Psychics: The Place of Mind in Nature." In this work, the terms "psychical monism" and "psychicalism" were proposed by Hartshorne to describe "panpsychism." Of the two proposals, it is "psychicalism" that has become widely accepted in philosophical jargon. Hartshorne intended "psychicalism" to be the opposite of the reductive physicalism that was popular at that time. In discussions about consciousness, he presented a third perspective, one that lies between reductive materialism and dualism.

[32] Griffin preferred the term "panexperientialism" over traditional terms such as "panpsychism" or "psychicalism." He justified this terminological proposal by saying that the term "psyche" suggests a higher form of experience (Griffin 1998, 78). Griffin defended a position he referred to as "panexperiential physicalism."

surface in Sprigge's philosophy, too, when he states that the whole consists of bits of experience. According to Sprigge, the whole of experience is constructed from these "bits of experience"— hence idealist panpsychism.

5. Panpsychism as a meta-view in the philosophy of mind: three crucial points

I have discussed the above stances within the framework provided by panpsychism as a meta-view. Consequently, it can be assumed that these positions propose a similar solution to the mind-body problem—namely, that at the fundamental level of reality, mental items exist. That these mental items are irreducible and exist at a fundamental level allows some proponents of panpsychism to argue that they can be causally effective. Within this framework, one can attempt to address issues (a) and (b) while simultaneously resolving the dilemma (Dil.). In reference to the frameworks of panpsychism as a meta-view, one can evaluate whether these positions provide an adequate response to the mind-body problem. Therefore, the proposal to solve the mind-body problem is itself neutral with respect to these positions.

Perhaps instead of elaborating these positions further and seeking answers in relation to naturalistic dualism, dual-aspect theory, non-standard physicalism, and idealism, it would be better to address the crucial problem for such a proposed solution to the mind-body problem: the combination problem.

It is important to comprehend how mental items, which are simple, singular, and distributed at the fundamental level of reality, can combine to produce the complex mental lives of human beings and other living organisms at a higher level of reality. The key to solving the mind-body problem lies in the combination problem. It can be proposed that the positions guided by the assumptions of panpsychism are merely labels that name the same approach to the mind-body problem but have the same metaphysical problem: the combination problem.

In this section, I will present three crucial points that follow from adopting panpsychism as a meta-view in the philosophy of mind.

(1) The solution to the mind-body problem is neutral. One can be a naturalistic dualist (property dualism), a proponent of the dual-aspect theory, a follower of non-standard physicalism, or an idealist and adopt a similar solution to the mind-body problem as outlined in section 3.1. Nevertheless, such a solution leads to the combination problem described in section 3.2., which also appears to be a problem shared by various positions in the philosophy of mind. This may indicate that these metaphysical positions are merely labels suggesting the same solution to the mind-body problem. Therefore, the solution is neutral concerning the different and seemingly contradictory positions.

(2) The combination problem is the most significant. Therefore, instead of focusing on solving the mind-body problem within a particular metaphysical system, we should rather focus on solving the combination problem. It is the most important metaphysical problem within panpsychism as a meta-view. Its resolution is crucial in solving the mind-body problem.

It is also important to highlight another crucial point of panpsychism as a meta-view: that the physical theories offer only a limited description of reality.[33] The adoption of M_1–M_4 in the previously discussed positions stems from this point. Without this point, Chalmers would not have formulated naturalistic dualism, Nagel would not have analyzed dual-aspect theory, Strawson would not have developed panpsychist physicalism, Sprigge would not have articulated idealistic panpsychism and would not have sought to define the true nature of reality.

Therefore, the proposed solution to the mind-body problem in section 3 must assume that mental items exist at the fundamental level of reality.

[33] Another issue to explore is that the arguments for adopting panpsychism that are commonly discussed in the literature are independent of the specific positions within the panpsychist framework. In other words, the arguments mentioned in this paper, such as the continuity argument (e.g., Chalmers, James), the anti-emergence argument (e.g., Strawson), or the intrinsic nature argument for panpsychism (e.g., Sprigge, Strawson) can be developed independently of the ontological and metaphysical position adopted within panpsychism. A valuable contribution to the literature on the types of argument for panpsychism is provided by Jarocki (2023).

(3) Physical theories offer a limited description of reality.[34] Thus, to resolve the mind-body problem, one must adopt the framework proposed by panpsychism as a meta-view, which assumes that mental items exist at the fundamental level of reality.

Consequently, it can be inferred that the described combination problem (2) also stems from point (3). If we believed that physics provides a comprehensive description of reality at the fundamental level, it would be unnecessary to discuss the combination problem, which arises from a bottom-up explanation.

5.1. Some critical remarks

One might ask whether I am treating panpsychism too broadly in this article. As mentioned in the introduction, panpsychism is typically understood as a metaphysical position in which mentality is considered fundamental and ubiquitous in the natural world. However, if this description is taken literally, it might imply that everything in existence, from atoms to the Taj Mahal, possesses a mind, consciousness, or experiences. However, contemporary panpsychists engaged in debates about the mind do not endorse such a view. Panpsychism understood in this way becomes caricatured.

Thus, in my proposal definition (D), I suggest separating the fundamental level, consisting of mental items, from which their appropriate composition allows for the production of conscious minds in human beings and other living organisms at a higher level. In this paper, I propose understanding panpsychism as a meta-view, emphasizing a bottom-up explanation and the MS. So, I believe we are faced with an alternative: either we consider panpsychism as a meta-view encompassing various positions, or we adopt a radical version of panpsychism in which everything that exists, from the Taj Mahal to atoms, possesses mental life.

On the other hand, in such a broadly understood panpsychism, the distinction between various positions in the metaphysics of mind may become

[34] The limitations of physics in explaining phenomenal consciousness are excellently described by Goff (2019). Historically, this issue was previously highlighted by Eddington (1928) and as mentioned in section 4.3 by Russell (1927).

blurred. As noted in section 5, these are perhaps not distinct metaphysical positions but merely labels suggesting the same solution to the mind-body problem (1). How I argue this problem can be addressed in a similar manner regardless of the adopted ontology and metaphysical positions, including naturalistic dualism, double aspect theory, physicalist panpsychism, or idealism. Therefore, instead of elaborating on various metaphysical systems, we should focus, as I suggest, on finding a comprehensive solution to the combination problem (2) (i)–(iii).

It is also pertinent to question whether panpsychism, as a meta-view, aligns with contemporary science. I have noted that within this meta-view, it is (3) proclaimed that physics has its limitations in describing reality. I will reference Hempel's dilemma,[35] within which we may ask, "Which physics?" (Hempel 1969). Are we addressing physical theories in their contemporary form or in a future form? Assuming the contemporary form, this stance is erroneous because contemporary physical theories might be proven false in several decades or centuries. On the other hand, in a future form, physical theories will be fundamentally different from today's and may encompass phenomena not assumed or explained by current physical theories. Consequently, it can be argued that panpsychism, as a meta-view, may not only provide a framework for understanding and potentially resolving the mind-body problem but also prove to be a framework for future science.

6. Conclusion

In the article, I offer an understanding of panpsychism as a meta-view in the philosophy of mind rather than as a position competing with others in the field. Using examples from the literature, I have discussed the positions of Chalmers' naturalistic dualism, Strawson's panpsychist physicalism, Sprigge's idealist panpsychism, and Nagel's remarks on the dual-aspect theory. Consequently, I suggest that these positions can all be accommodated within the panpsychism meta-view, as panpsychism does not inherently contradict any of them. Thus, these positions may be explored within

[35] Accordance to the Hempel's dilemma, the thesis of physicalism is either false or empty.

the framework of panpsychism as a meta-view in addressing the mind-body problem. Therefore, I have indicated that the positions discussed in section 4 should represent different labels for a similar solution to the mind-body problem. Hence, rather than elaborating on distinct metaphysical positions, it may be more beneficial to focus on a comprehensive solution to the combination problem (2) (i)–(iii), as this is the problem that impedes adequate resolution of the mind-body problem. I have also pointed out that within the framework of panpsychism as a meta-view, it should be maintained that physics has its limitations and does not adequately describe reality.

Acknowledgements

The author would like to thank the anonymous reviewer, as well as Tadeusz Szubka and Karol Polcyn, for their comments and suggestions.

References

Blamauer, Michael, ed. 2011a. *The Mental as Fundamental: New Perspectives on Panpsychism*. Frankfurt: Ontos-Verlag. https://doi.org/10.1515/9783110319859

Blamauer, Michael. 2011b. "Taking the Hard Problem of Consciousness Seriously: Dualism, Panpsychism, and the Origin of the Combination Problem." In *The Mental as Fundamental: New Perspectives on Panpsychism*, edited by Michael Blamauer, 99–116. Frankfurt: Ontos-Verlag. https://doi.org/10.1515/9783110319859.99

Bradley, Francis Herbert. 1893 [1969]. *Appearance and Reality*. Oxford: Oxford University Press.

Brüntrup Godehard and Ludwig Jaskolla. 2017a. "Introduction." In *Panpsychism: Contemporary Perspectives*, edited by Godehard Brüntrup and Ludwig Jaskolla, 1–16. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199359943.003.0001

Brüntrup Godehard and Ludwig Jaskolla, eds. 2017b. *Panpsychism: Contemporary Perspectives*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199359943.001.0001

Chalmers, David. 1995. "Facing Up to the Problem of Consciousness." *Journal of Consciousness Studies*, 2(3), 200–19. https://doi.org/10.1093/acprof:oso/9780195311105.003.0001

Chalmers, David. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.

Chalmers, David. 2003. "Consciousness and its Place in Nature." In *Blackwell Guide to the Philosophy of Mind*, edited by Stephen P. Stich and Ted. A. Warfield, 102–42. Oxford: Blackwell. https://doi.org/10.1002/9780470998762.ch5

Chalmers, David. 2017a. "The Combination Problem for Panpsychism." In *Panpsychism: Contemporary Perspectives*, edited by Godehard Brüntrup and Ludwig Jaskolla, 179–214. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199359943.003.0008

Chalmers, David. 2017b. "Panpsychism and Panprotopsychism." In *Panpsychism: Contemporary Perspectives*, edited by Godehard Brüntrup and Ludwig Jaskolla, 19–47. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199359943.003.0002

Cobb, John B. 2008. *Whitehead Word Book: A Glossary With Alphabetical Index to Technical Terms in Process and Reality*. Claremont: P&F Press.

Coleman, Sam. 2006. "Being Realistic: Why Physicalism May Entail Panexperientialism." *Journal of Consciousness Studies*, 13(10-11), 40–52.

Coleman, Sam. 2012. "Mental Chemistry: Combination for Panpsychists." *Dialectica*, 66(1), 137-166. https://doi.org/10.1111/j.1746-8361.2012.01293.x

Coleman, Sam. 2014. "The Real Combination Problem: Panpsychism, Micro-subjects, and Emergence." *Erkenntnis*, 79, 19–44. https://doi.org/10.1007/s10670-013-9431-x

Coleman, Sam. 2017. "Panpsychism and Neutral Monism: How to Make up One's Mind." In *Panpsychism: Contemporary Perspectives*, edited by Godehard Brüntrup and Ludwig Jaskolla, 249–82. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199359943.003.0011

Crane, Tim. 2000. "The Origins of Qualia." In *The History of the Mind-Body Problem*, edited by Tim Crane and Sarah Patterson, 169–94. London: Routledge. https://doi.org/110.4324/9780203471029-12

Davidson, Donald. 1970. "Mental Events." In *Experience and Theory*, edited by Lawrence Foster and J. W. Swanson, 207-224. Amherst, Massachusetts: University of Massachusetts Press. https://doi.org/10.1093/0199246270.003.0011

Eddington, Arthur Stanley. 1928. *The Nature of the Physical World*. Cambridge: Cambridge University Press.

Feigl, Herbet. 1958. *The "Mental" and the "Physical:" The Essay and a Postscript*. Minnesota: University of Minnesota Press.

Goff, Philip. 2009a. "Can the Panpsychist get Round the Combination Problem?" In *Mind that Abides: Panpsychism in the New Millennium*, edited by David Skrbina, 283–304. Amsterdam: John Benjamins. https://doi.org/10.1075/aicr.75.10gof

Goff, Philip. 2009b. "Why Panpsychism Doesn't Help us Explain Consciousness." *Dialectica*, 63(3), 289–311. https://doi.org/10.1111/j.1746-8361.2009.01196.x

Goff, Philip. 2017a. *Consciousness and Fundamental Reality*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780190677015.001.0001

Goff, Philip 2017b. "The Case for Panpsychism." *Philosophy Now*, 121, 6–8.

Goff, Philip. 2017c. "The Phenomenal Bonding Solution to the Combination Problem." In *Panpsychism: Contemporary Perspectives*, edited by Godehard Brüntrup and Ludwig Jaskolla, 283–304. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199359943.003.0012

Goff, Philip. 2019. *Galileo's Error*. New York: Pantheon.

Goff, Philip and Alex Moran, eds. 2022. *Is Consciousness Everywhere? Essays on Panpsychism*. Exeter: Imprint Academic. https://doi.org/10.53765/20512201.28.9.009.

Goff, Philip, Seager William and Sean Allen-Hermanson. 2022. "Panpsychism." In *The Stanford encyclopedia of philosophy* (Summer 2022 Edition), edited by Edward N. Zalta. Last updated May 13, 2022. https://plato.stanford.edu/entries/panpsychism/

Griffin, David Ray. 1997. "Panexperiential Physicalism and the Mind-Body Problem." *Journal of Consciousness Studies*, 4(3), 248–68.

Griffin, David Ray. 1998. *Unsnarling the World-Knot*. Eugene: Wipf and Stock.

Hartshorne, Charles. 1950. "Panpsychism." In *A History of Philosophical Systems*, edited by Vergilius Ferm, 442–53. New York: The Philosophical Library.

Hartshorne, Charles. 1977. "Physics and Psychics: The Place of Mind in Nature." In *Mind and Nature*, edited by John B. Cobb and David Ray Griffin, 90–122. Lanham: University Press of America.

Hempel, Carl Gustav. 1969. "Reduction: Ontological and Linguistic Facets." In *Philosophy, Science and Method: Essays in Honor of Ernest Nagel*, edited by Sidney Morgenbesser, Patrick Suppes and Morton White, 190-197. London: St. Martin's Press.

Jackson, Frank. 1986. "What Mary Didn't Know". *The Journal of Philosophy*, 83(5), 291–5. https://doi.org/10.2307/2026143

James, William. 1890. *The Principles of Psychology*. New York: Henry Holt and Company. http://dx.doi.org/10.1037/11059-000

Jarocki, Jacek. 2023. Three Kinds of Arguments for Panpsychism. *Acta Analytica*, 39(3). https://doi.org/10.1007/s12136-023-00566-z

Kim, Jaegwon. 1993. *Supervenience and Mind: Selected Philosophical Essays*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511625220

Kim, Jaegwon. 1998. *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. Cambridge: MIT Press. https://doi.org/10.7551/mitpress/4629.001.0001

Kripke, Saul Aron. 1980. *Naming and Necessity*. Cambridge, Massachusetts: Harvard University Press. http://dx.doi.org/10.1007/978-94-010-2557-7_26

Levin, Janet. 2022. *The Metaphysics of Mind*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108946803

McHenry, Leemon. 2010. "Sprigge's Ontology of Consciousness." *Royal Institute of Philosophy Supplement*, 67, 5–20. https://doi.org/10.1017/S1358246110000093

Nagel, Thomas. 1974. "What Is it Like to Be a Bat?" *Philosophical Review*, 83, 435–50. https://doi.org/10.2307/2183914

Nagel, Thomas. 1979. "Panpsychism." In *Mortal Questions*, edited by Thomas Nagel, 181-195. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781107341050.015

Nagel, Thomas. 1986. *The View from Nowhere*. Oxford: Oxford University Press. https://doi.org/10.2307/2220404

Nagel, Thomas. 2012. *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature is Almost Certainly False*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199919758.001.0001

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

Roelofs, Luke. 2020. "Can We Sum Subjects? Evaluating Panpsychism's Hard Problem." In *The Routledge Handbook of Panpsychism*, edited be William Seager, 245–58. New York: Routledge.

Rosenberg, Greg. 1996. *Consciousness and Causation: Clues Toward a Double-Aspect Theory*. Bloomington, Indiana: Indiana University.

Rosenberg, Greg. 2004. *A Place for Consciousness: Probing the Deep Structure of the Natural World*. Oxford: Oxford University Press.

Rosenberg, Greg. 2017. "Land Ho? We are Close to a Synoptic Understanding of Consciousness." In *Panpsychism*, edited by Godehard Brüntrup and Ludwig Jaskolla, 130–53. Oxford: Oxford University Press.

Russell, Bertrand. 1927. *The Analysis of Matter*. London: Kegan Paul.

Seager, William. 1995. "Consciousness, Information and Panpsychism." *Journal of Consciousness Studies*, 2, 272–88.

Seager, William. 2010. "Panpsychism, Aggregation and Combinatorial Infusion." *Mind and Matter*, 8(2), 167–84.

Seager, William. 2017. "Panpsychist Infusion." In *Panpsychism*, edited by Godehard Brüntrup and Ludwig Jaskolla, 229-248. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199359943.003.0010

Seager, William, ed. 2020. *The Routledge Handbook of Panpsychism*. London and New York: Routledge. https://doi.org/10.4324/9781315717708

Searle, John Rogers. 2005. *Mind: A Brief Introduction*. New York: Oxford University Press.

Shannon, Claude E. 1948. "A Mathematical Theory of Communication." *Bell Systems Technical Journal*, 27: 379–23. https://doi.org/10.1002/j.1538-7305.1948.tb00917.x

Skrbina, David. 2017. *Panpsychism in the West*. Cambridge, Massachusetts: The MIT Press.

Smart, J. J. C. 1959. "Sensations and Brain Processes." *The Philosophical Review*, 68(2), 141-156. https://doi.org/10.2307/2182164

Sprigge, Timothy. 1983. *The Vindication of Absolute Idealism*. Edinburgh: Edinburgh University Press. https://doi.org/10.1515/9781474472814

Stoljar, David. 2001. "Two Conceptions of the Physical." *Philosophy and Phenomenological Research*, 62(2), 253–81. https://doi.org/10.2307/2653699

Strawson, Galen. 2008a [2003]. "Realistic Materialism." In *Real Materialism: And Other Essays*, edited be Galen Strawson, 19–52. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199267422.003.0002

Strawson, Galen. 2008b. "Realistic Monism: Why Physicalism Entails Panpsychism." In *Real Materialism: And Other Essays*, edited be Galen Strawson, 55–74. Oxford: Oxford University Pres. https://doi.org/10.1093/acprof:oso/9780199267422.003.0003

Strawson, Galen. 2015. "Real Materialism (with New Postscript)." In *Consciousness in the Physical Word*, edited by Torin Alter and Yuijn Nagasawa, 161-208. Oxford: Oxford University Press.

Strawson, Galen. 2017a. "Physicalist Panpsychism." In *The Blackwell Companion to Consciousness*, edited by Susan Schneider and Max Velmans, 374–90. Oxford: Blackwell. https://doi.org/10.1002/9781119132363.ch27

Strawson, Galen. 2017b. "Mind and Being: The Primacy of Panpsychism." In *Panpsychism*, edited by Godehard Brüntrup and Ludwig Jaskolla, 75–112. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199359943.003.0004

Strawson, Galen. 2020. "What Does 'Physical' Mean? A Prolegomenon to Physicalist Panpsychism." In *The Routledge Handbook of Panpsychism*, edited by William Seager, 317–99. London and New York: Routledge.

Westphal, Jonathan. 2016. *The Mind-Body Problem*. Cambridge, Massachusetts: The MIT Press. https://doi.org/10.7551/mitpress/10776.001.0001

Whitehead, Alfred North. 1925. *Science and the Modern World*. Cambridge: Cambridge University Press.

Organon F 31 (3) 2024: 309–336
https://doi.org/10.31577/orgf.2024.31305

ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

Essence and Modality: Continued Debate

Andrew Dennis Bassford*

Received: 29 January 2024/ Accepted: 22 July 2024

Abstract: Here I offer a critical evaluation of modalism about essential properties. To that effect, I begin by rehearsing Fine's now infamous counterexamples to pure modalism. I then consider two recent defenses of it, offered by Livingstone-Banks and Cowling, respectively. I argue that both defenses fail. Next I consider the most plausible variety of impure modalism – sparse modalism – which has recently been defended by Wildman and de Melo. Skiles has argued that sparse modalism fails too. I argue that Skiles's counterexamples misfire; nonetheless, his conclusion that, like pure modalism, sparse modalism is too broad, is on the right track. And so, I offer an original objection – the sparse modal propria counterexample – to show that this is so. I conclude by considering ways the modalist might once again modify her account to circumvent this new objection and improve the account's extensional adequacy.

Keywords: Essence and Essentialism; Essentialism and Quantified Modal Logic; De Re Modality; Properties; Kit Fine.

* University of Texas at Austin
https://orcid.org/0000-0002-7118-4236
University of Texas, 2210 Speedway, Stop C3500, Austin, Texas, USA 78712
a.d.bassford@utexas.edu

1. Pure modalism

What is an essential property? More to the point, do essential properties admit of a reductive analysis or not? And if they do, how ought we to analyze what it is to be an essential property? Within the dispute over the character of essential properties, answers to the second question divide reductivists about essences from nonreductivists (or primitivists); and answers to the third distinguish reductivists from one another.

"Modalism" denotes a family of reductive analyses. The unifying thought behind all varieties of modalism is that essences are fundamentally modal in character. Modalism comes in two varieties. Pure modalists suppose that essences are completely reducible to de re modality. Impure (or hybrid) modalists suppose only that essences are partly reducible to it; some other nonmodal phenomenon is appealed to in the analysans as well.

The most popular (and plausible) version of pure modalism is the Existential-Modal Account (EMA). According to EMA, Socrates is essentially human iff: necessarily, Socrates is human if he exists. Put formally, $\Box(\exists x(x = s) \rightarrow H(s))$. Or, put in terms of possible worlds, for all worlds, w, if Socrates exists at w, then Socrates is human at w too.

Modalism (and especially EMA) enjoyed almost unanimous acceptance among essentialists between the 1950s and 1990s, and especially during the 1970s and '80s, by and large as a result of the work of Kripke (1972/1980) and Plantinga (1970, 1979), among others.[1] It captures well the intuition that essential properties of an object are necessary for the existence of their bearers. Accordingly, it also captures well the intuition that paradigmatic accidental properties are not necessary for their bearers, but rather are possibly had or not had.[2] On this assumption, essentialists defined themselves by a commitment to this thesis; and antiessentialists, like Stalnaker (1979), by an opposition to it.[3]

---

[1] Plantinga (1970, 474): "[A] property is essential to Socrates just in case he has it and there is no world in which he has its complement…"

[2] Aristotle (c. 350 BC[b], Topics 102b5-10): "An accident is… something which may possibly either belong or not belong to any one and the self-same thing."

[3] Stalnaker (1979, 344): "[F]or every individual and every property, there are possible worlds in which the individual has the property and possible worlds in which it does not."

Modalism has since fallen is disrepute. In 1994, Kit Fine famously offered a series of counterexamples to EMA, intended to show that, even if it captures an important necessary condition for being an essential property, as it stands, the analysis is insufficient. As Roca-Royes (2011: 66) notes, "Fine's counterexamples to EMA have been widely accepted; to a degree that is unusual for philosophical arguments."

Fine offers five counterexamples to EMA. His first is that EMA would count existence as an essential property of Socrates, since necessarily, he has the property of existing if he exists. But, intuitively, neither Socrates nor any other mere contingent being has the property of existing essentially. Maybe God exists essentially, but Socrates does not. His second is that EMA would count distinctiveness properties as essential, since necessarily, Socrates has the property of being distinct from the Eiffel Tower if he exists. "But it is not essential to Socrates that he be distinct from the Tower, for there is nothing in his nature which connects him in any special way to it" (Fine 1994: 5). His third is that EMA would count necessary truths as essential, since necessarily, Socrates has the property of being such that $2 + 2 = 4$ if he exists. But it is no part of Socrates's essence that some necessary truth obtains. His fourth is that EMA would count metaessential properties (i.e., essential properties of other objects) as essential, since necessarily, Socrates has the property of being such that being human is essential to Plato, if Socrates exists. But, if this is so, then – "O happy metaphysician!" (Fine 1994: 6) – it would follow that if one discovers the essence of one thing, one also thereby discovers the essence of everything else. But this is intuitively false.

And his fifth, and perhaps most famous, objection is that EMA would count set-membership properties as essential. Necessarily, Socrates has the property of being the sole member of the singleton set {Socrates} if he exists. "But, intuitively, this is not so. It is no part of the essence of Socrates to belong to the singleton" (Fine 1994, 5). And so, for not just one, but five reasons, EMA is too broad and so extensionally inadequate.

Fine's diagnosis is that EMA (and pure modalism generally) has failed in each case because any adequate analysis of essential properties must answer the question, asked of any particular thing 'What is it?' Not every property that is necessary for the existence of its bearer answers this question, and so

pure modalism goes sideways. Fine recommends we adopt an alternative non-reductivist position, the Real Definitional Account (RDA) whereby essences are treated as a primitive (perhaps hyperintensional) notion.

2. Thesis

The goal for this paper is to critically evaluate several recent modalist responses to Fine's counterexamples. Modalists have responded to Fine in one of two ways. Some have argued that Fine's counterexamples to EMA are not genuine, for one reason or another. This amounts to a defense of pure modalism. Other modalists have gone impure. They argue that Fine's counterexamples only show that some distinction should be drawn between kinds of properties an object might possess. Most promisingly, some have followed Armstrong (1979) and Lewis (1983a) in drawing a sparse vs. abundant property distinction. Using this resource, they then show that at least one form of modalism (viz., sparse modalism) can escape his counterexamples unscathed.

I'll argue for two theses here. The first is that two of the best pure modalist defenses fail. Fine's counterexamples are genuine, and so the modalist's best bet is to go impure. And the second is that the most plausible variety of impure modalism – sparse modalism – fails too. Like pure modalism, it is extensionally inadequate because too broad. I'll offer an original counterexample to this effect. After offering my counterexample, I'll consider and refute one potential response to it before offering my diagnosis of why modalism (of both the pure and impure varieties) is ultimately unsatisfying.

3. Livingstone-Banks's defense

The pure modalist responds to Fine by denying that his counterexamples are genuine. Both Livingstone-Banks (2017) and Cowling (2013) have recently taken this approach. I take their responses to be representative of the best pure modalist responses available in the dialectic. Nonetheless, they are both problematic.

Livingstone-Banks's basic move is to challenge the appeals to intuition (either implicitly or explicitly) embedded in the statements of all five cases. He says that Fine faces a dilemma:

> There are two ways we might interpret these intuitions. On the one hand, they might be pre-philosophical opinions. This raises the epistemic question of how we might have access to essentialist facts sufficient for our intuitions to be an accurate guide to essence, and the methodological question of how we can insist that any theory of essence be beholden to those intuitions (especially before we have a satisfactory answer to the epistemic question). On the other hand, we might understand those intuitions as being philosophical in nature, in which case the examples Fine presents in his critique of [modalism] should not be interpreted as arguments against [EMA], for to do so is merely to beg the question in favor of [RDA]… (Livingstone-Banks 2017, 816–18)

And so, Fine's counterexamples either beg the question in assuming his own position (RDA), or they rely on purportedly clear pre-theoretical intuitions about essences – but there are none.

On its face, this is a smart move. Nonetheless, there are several ways Fine might sail between the two horns. Most promisingly, I think the critic of modalism should follow Lewis (in this dialectical context) in supposing that:

> Our 'intuitions' are simply opinions; our philosophical theories are the same. Some are commonsensical, some are sophisticated; some are particular, some general; some are more firmly held, some less. But they are all opinions, and a reasonable goal for a philosopher is to bring them into equilibrium. (Lewis 1983b, x)

Like Livingstone-Banks, Lewis endorses a kind of Prior Opinions, or Common Law, account of intuitions. He then makes four distinctions. First, he distinguishes between sources of prior opinions. Following him, we can say that there are (i) our own opinions, (ii) commonsense opinions (which may) include the opinions of scientists in our community), and (iii) the opinions of other philosophers; relevant for present purposes, the opinions of other metaphysicians of essence. Second, he distinguishes between beliefs

firmly held and more loosely held. Third, he distinguishes between general intuitions and particular intuitions; let's call an intuition general if it is about what essential properties are generally; we'll call it particular if pertains to this or that particular case, whether this or that property is essential to a certain object or not. And finally, he says that the job of the philosopher is to bring beliefs into equilibrium. Equilibrium is a matter of degree. Nonetheless, the more a belief coheres will all three sets of beliefs, especially those firmly held, the better an equilibrium it strikes (i.e., the more intuitive it is, all things considered).

Using Lewis's distinctions, we can then show that Livingstone-Bank's disjunctive premise is false. He says that Fine is either begging the question in assuming RDA, or he is relying on nonexistent clear pre-theoretical intuitions. This isn't so. First, Fine is correct in saying that our general intuitions about essences is that they should answer the 'What is it?' type questions asked of some object. This coheres well with all three sets of prior opinion, whether one finds RDA plausible or not. But second, Fine need not appeal to this general intuition at all in order for his counterexamples to go through. EMA would need to count the property of being such that $2 + 2 = 4$ as essential to Socrates, as such. But, our particular intuitions say that this isn't so. And so, EMA's verdict on this case fails to strike a decent equilibrium with our beliefs about this case, as well as all of the other cases too. His counterexamples are successful because they show that EMA is out of step with our particular intuitions, not our more general, and our intuitions about the cases are fairly clear.

So, Livingstone-Banks's defense of pure modalism fails. There is no dilemma here. The way to strike the best equilibrium is to reject EMA.

4. Cowling's defense

Cowling has offered a second defense of pure modalism. Cowling's basic move is to challenge our previous claim, that the best way to strike an equilibrium with our prior opinions is to reject EMA. He argues that a distinction ought to be made between essences, on the one hand, and natures, on the other. He says that Fine's counterexample to EMA only seem genuine because we have mistakenly identified the two. Pure modalism is

the correct analysis of essence. Fine's counterexamples only show that it's the wrong analysis of natures.

Cowling's main argument for this distinction appeals to the need to preserve the integrity of the debate about the scope of essence. He is particularly concerned to find a place in the debate for Quine, who was famously an eliminativist about essences and essential properties. Quine supposed that there are no essential properties and, more pointedly, that de re modality is incoherent. But, he did suppose that objects have natures, and he set about working to discover what the natures of various objects are. In doing so, he thought he was addressing the question 'What is it?' And so, we face (another) dilemma. Either Quine was conceptually confused, which then ruins the integrity of his contribution to the discussion; or we must say that Fine is conceptually confused, and so his counterexamples misfire and target the wrong thesis. Cowling supposes we should preserve the integrity of the debate on the scope of essential properties (involving Quine), and so reject that Fine's counterexamples to EMA are decisive. The way to strike the best equilibrium is to reject that essences are identical to natures; this is the important lesson we should learn from Fine's cases.

On its face, this is a smart move too. However, the defender of Fine should say two things here. First, she should concede that preserving a place in the dispute for Quine would strike the optimal equilibrium, all the while denying that this is a sufficient reason for making an (apparently) artificial distinction. Returning to a previous point, intuitively, the two terms 'essence' and 'nature' are synonymous. I myself make no distinction between the two. Commonsense likewise makes no distinction between the two. And importantly, many other metaphysicians (both contemporary and historical) do not make a distinction between the two either.

Cowling says that the essence of a thing answers what is possible or impossible for it, whereas the nature of a thing answers the 'What is it' question. But consider Descartes (1647) on this point: "The impossibility of existing without a valley is part of the nature of a mountain; and it belongs just as much to the nature of the human mind that it is what it is…" (AT VIIIB 348). He evidently thinks the nature of a thing answers both questions. Or, if a distinction is made, it is made as Aristotle (c. 350 BC[a]) put it in his philosophical lexicon: "'Nature' means the essence of natural

objects… By an extension of meaning from this sense of 'nature' every essence in general has come to be called a 'nature,' because the nature of a thing is one kind of essence" (Meta. Δ4 1014b35-1015a13). Incidentally, in following Cowling, we would keep Quine in the dispute only at the expense of excluding the two of them. This is an even less happy equilibrium than the alternative.

And a second problem with Cowling's response is that, as he himself states the objection, "[T]he modal view of essence, once separated from the modal view of nature, is theoretically uninteresting and irrelevant for metaphysical inquiry. But since essence cannot plausibly be viewed as uninteresting or irrelevant, the essence-nature distinction must be rejected" (2013: 11). In response to this objection, Cowling maintains that there does remain one theoretically interesting and relevant role for essence to play in our metaphysical inquiry. Namely, we can ask what properties an object has that are both essential to it and of its nature, and that this inquiry will tell us something very interesting and relevant about the thing (i.e. its natural essence).

But this response is insufficient too. This would make essences theoretically interesting and relevant only in some derivative or parasitic sense. I, the ordinary person, and most other metaphysicians suppose that knowing the essence of something – quite apart from knowing anything else about the object – is intrinsically valuable and interesting. Essences inherently carry with them a kind of metaphysical gravitas, as Wildman puts it. In order to preserve this intuition, Cowling would need to show how knowing the essence, once divorced from nature, is by itself valuable. For my part, I do not see how this challenge can be met if EMA is assumed.

And so, for these reasons, Cowling's defense of pure modalism is unsuccessful. It is better not to make a distinction between essences and natures; and so, it is better to suppose that Fine's objection are not misdirected at a modal view of natures, but rather are adequately directed at the modal view of essence.

5. Sparse properties

Other pure modalist responses are possible; nonetheless, I think these two are representative of the best responses available in this dialectic. And

so, since they fail, the modalist should accept Fine's critique of pure modalism and work instead to add some further stipulation into their analysis to block his counterexamples.

This is the route that most modalists have gone. Impure modalists reject Fine's conclusion that modalism, in general, is false. Rather, they maintain that his cases just show "the more modest conclusion that some naturalness restriction needs to be imposed on candidate properties for essentiality" (Vetter 2009: 3). This "naturalness" restriction is drawn in various ways (cf. Robertson & Atkins 2018: §2). At present, sparse modalism is the most popular (and plausible) variety of impure modalism (cf. Wildman 2016). It carves out the distinction in terms of sparsity.

Sparse properties are properties that carve up nature at its joints. This is usually cashed out by distinguishing properties that (a) ground qualitative similarity and dissimilarity between objects; (b) track the causal powers of objects; and (c) do, or might, figure into our best sciences of the laws of nature – from those properties that do not. Whereas members of the first class are fundamental and genuinely seem to pick out natural features of the world, members of the second are derivative and correspond to any property that can be thought whatsoever; for any two objects, there will correspond a property that picks them out. Examples of sparse properties (might) include: being green; being an electron; and being a planet. Examples of nonsparse (i.e. abundant) properties include: being grue; being an electron that comprises my left ear; and being a planet or not a planet. We might need to appeal to abundant properties in our best formal semantics; but otherwise they form a metaphysically useless and redundant class.

The distinction between sparse vs. abundant properties was originally introduced by Armstrong (1979), and then further championed by Lewis (1983a). According to Lewis, an appeal to sparse properties is needed in our best systematic metaphysics. (Note that both Armstrong and Lewis appeal to sparse properties independently of discussion about the character of essences, and so, by appealing to sparse properties, the sparse modalist can quickly escape the charge that their modified account is merely ad hoc). Lewis provided something like a laundry list of tasks for which sparse properties are needed. As Wildman recounts, sparse properties

> could be employed in analyzing the laws of nature, causation, intrinsicality, and supervenience; are needed to account for Moorean facts of common sense, provide a minimal notion of physicalism, handle Kripke's rule-following worries, and respond to Putnam's objections to metaphysical realism; and, perhaps most importantly, are required to determine what, in the fundamental sense, objects are like – without the elite properties fixing the objective facts about the comparative characters of objects, we'd have to fall back into the unpalatable position that the only real structure of the universe is its cardinality. (Wildman 2013, 763)

That they might play some of these roles in our best systematic metaphysics gives us one reason to make the sparse vs. abundant property distinction and suppose that there are sparse properties.

As noted by Schaffer (2004), at least two conceptions of sparse properties are possible, depending on how one understands criterion (c), what counts as our "best sciences." On a fundamental conception of sparsity, only those properties that are needed in fundamental physics are sparse (i.e., only those properties needed to describe the microphysical world, from which everything else is built up). On the assumption that scientists are only concerned with sparse properties, this would then amount to that view that all of science is either physics or stamp-collecting. And on a broadly scientific conception of sparsity, properties that are appealed to in our softer sciences (e.g., biology, psychology, etc.) also count as sparse. Schaffer has argued that, of the two, the latter is the best, and most sparse modalists have followed him in this respect. On this latter conception, a property like being human would then count as sparse, since it is a property appealed to in our biological sciences.

6. (Absolute) Sparse Modalism

So understood, according to the Sparse Modal Account (SMA), Socrates is essentially human iff (1), necessary, Socrates is human if he exists; and (2) being human is sparse. Put formally, where 'S' designates the second-order property of being sparse: $\Box(\exists x(x = s) \rightarrow (H(s) \wedge S(H)))$. Or, put in

terms of possible worlds, for all worlds, w, if Socrates exists at w, then he is human at w too, and being human is a sparse property (across all worlds).

SMA has several virtues. First, like EMA, it makes sense of our intuition that essential properties are those properties without which their bearer could not possibly be. Moreover, SMA also satisfies our intuition that essential properties are intrinsically valuable and interesting in their own right. Knowing that Socrates is essentially human, then, tells us something about how he fits into the joints of nature, how he is objectively characterized and is similar to other kinds of natural things, and, given what kind of thing he is, into which scientific domain he falls and to which science it belongs to study him.[4] Notice also that, because sparse properties genuinely characterize things and ground objective relations of qualitative similarity between things, SMA also apparently captures Fine's important general intuition as well – that the essence of a thing ought to (at least partly) answer the 'What is it?' type questions about that thing.

Moreover, by adding in a sparsity condition to EMA, SMA can easily circumvent at least four of Fine's original counterexamples. SMA need not count being such that $2 + 2 = 4$ as essential to Socrates because this property is not sparse. It grounds no qualitative similarity; it tracks no causal powers; and it isn't needed outside of formal semantics. SMA need not count being such that Plato is essential human, likewise, because it is not sparse. Plato's being essential human grounds no qualitative similarity between Socrates and anything else; it tracks none of his causal powers; and no science needed to described Socrates's place in the world need appeal to it to characterize him. And, in a similar vein, SMA need not count the property of existing as essential to Socrates either. Several arguments might be offered to this effect. Here's one: An existing Socrates is qualitatively identical to a non-existing Socrates. The property of existing makes no qualitative difference to him (even if it makes a qualitative difference to the

[4] Note that Cowling (2013) supposed that the nature of a thing is best characterized by appealing to its sparse properties. Perhaps for this reason, some suppose that Cowling is also a sparse modalist, and it is not uncommon to see Wildman's and Cowling's defenses lumped together in the literature.

world). And so, the property of existing is not sparse, and so not essential to Socrates, as such.[5]

Fine's other two counterexamples gesture at relations between, in this case, Socrates and something else. SMA says nothing about relations, only about properties. Nonetheless, one sparse modalist response is to allow that there are sparse relations and that those relations might count as essential, and to do so in such a way that distinctiveness properties and membership properties are excluded while (say) a thing's origins are not. This is the route Wildman (2013) takes. In this way, SMA might be said to consist of two theses: sparse property modalism (SPM) and sparse relation modalism (SRM):

> [SPM] Object x essentially has property Φ iff (1) necessarily, x has Φ if it exists; and (2) Φ is a sparse property.

> [SRM] Object x essentially bears relation Ψ to y iff (1) necessarily, if x exists, then Ψ holds of the ordered pair <x, y>; and (2) Ψ is a sparse relation.

With SRM in place, the sparse modalist can then offer two responses to the case of distinctiveness properties. First, she can deny that Socrates's being distinct from the Eiffel Tower is necessarily true if Socrates exist, since, understood in this way, there will be worlds where Socrates exists and the Eiffel Tower does not (i.e., the ordered pair <Socrates, Eiffel Tower> will fail to hold at any world at which one of the relata do not exist). And so, distinctiveness properties fail criterion (1) of SRM. And second, she can say that distinctiveness properties are not sparse, and so fail criterion (2) of SRM as well. This is because distinctiveness properties are not needed in our best sciences to completely describe the world without redundancy; nor do they track causal powers of their bearers.

[5] Additionally, the modalist can point out that in saying God essentially exists, all we really mean here is that God is a necessary being; i.e. □(∃x (x = God)). But this, of course, is not what we're saying in concluding that Socrates essentially exists, but only that necessary, he exists if he exists. The paradox here is merely apparent.

7. Socrates and {Socrates}

This just leaves Fine's most infamous case – Socrates's being the sole member of {Socrates}. This is his most famous because, unlike distinctness properties, there's supposed to be a kind of essential asymmetry involved here (Dunn 1990: 77, 89). Presumably, we want to say Socrates is not essentially a member of {Socrates}; but we do want to say that {Socrates}'s having Socrates as a member is essential to it. With the resources of only extensional and intensional analysis, both claims would amount to the same thing. The relation is either sparse or it isn't.

And so, the modalist must make one of two choices here.[6] On the one hand, she can deny that sets have essences and then dispel the air of paradox in saying that {Socrates} does not essentially have Socrates as a member, that way she can deny that Socrates is essentially a member of {Socrates}. On the other hand, she can affirm that sets might have essences and then dispel the air of paradox in saying that Socrates is essentially a member of {Socrates}, that way she can affirm that {Socrates} is essentially membered by Socrates. In either case, she would need to offer us some powerful error theory to explain away the apparent asymmetry we detect in the case.

Neither response is perfectly satisfying. Nonetheless, I think the sparse modalist would do best to deny that membership properties are sparse since Socrates's being a member of {Socrates} tracks none of his causal powers or the like at all. This would entail that neither does {Socrates} essentially have Socrates as a member. On this point, Skiles (2015: 7) has said: "If the sparse modalist is unable to capture the basic essentialist intuition that it be essential to {Socrates} to have Socrates as a member, then the sparse modalist seems unsuccessful at stating what is necessary for essentiality." But this intuition is defeasible. The sparse modalist can offer a more or less plausible three-part error theory on this subject.

First, she can appeal to Cowling's intuition, discussed earlier. Cowling said that a distinction should be made between a thing's nature and its essence. This distinction is decidedly artificial; 'nature' and 'essence' are

[6] Strictly speaking, this dilemma is false. The sparse modalist could somehow modify her account to make room for essential asymmetry. This point will be discussed at some later point, below.

semantically (and metaphysically) interchangeable. Nonetheless, even if there is no difference in the denotation between the two terms, we might agree with Cowling that there is a difference in the connotation between the two terms. It seems right to say that sets have essences. But, the sparse modalist can ask rhetorically, does it also seem right to say that sets have natures? Put this way, I think the answer is no. Only natural things have natures, and sets do not seem to be natural things.

Second, the sparse modalist can point out that, in saying that it is not the case that {Socrates} essentially has Socrates as a member, we are not thereby committed to saying that {Socrates} only accidentally has Socrates as a member, understood in the usual sense. The sparse modalist can appeal to the threefold distinction between essences, accidents, and necessary-for-existence properties. We can still grant that it is necessary for the existence of {Socrates} that it have Socrates as a member; it is impossible for it to exist as the selfsame thing unless it meets this condition. But this does not mean that it is essential to it that it have this property. We're only tracking the fact that it is impossible for it to be otherwise. But this, as Fine has taught us, is not equivalent to its having some essence.

And third, and perhaps most radically, the sparse modalist might attempt to accommodate our intuitions of asymmetry in the Socrates / {Socrates} case by conceding that some asymmetry exists, but denying that it is an essential asymmetry. Rather, it exists only in some cognate notion to essentiality. There exists here instead an analytic asymmetry.

Analyticity may be defined in several different ways, but consider Kant's (1781) definition that a proposition of the form 'A is B' is analytic iff "the predicate B belongs to the subject A as something which is (covertly) contained in this concept A… If I say, for instance, 'All bodies are extended,' this is an analytic judgment. For I do not require to go beyond the conception which I connect with 'body' in order to find extension as bound up with it" (Intro. IV, p. 48). The modalist can then say, even though the membership relation is essential to neither of the relata, nonetheless there's an analytic relation between '{Socrates}' and its property of having Socrates as a member; whereas there is not an analytic relation between 'Socrates' and his being a member of {Socrates}. The concept '{Socrates}' is defined by its relation to Socrates, but 'Socrates' is in no way defined by its relation

to {Socrates}. Because analyticity is in some respects very similar to essentiality, it's reasonable that we might have confused one for the other.

Some will likely want to bite back against this third move. What sense can be made in saying some judgment is analytic, and yet the property expressed by the judgement's predicate is not essential to its subject? Kripke (1972/1980) has shown us that some property might essentially belong to an object without the corresponding judgement being analytic (e.g., 'Water is H20); but an analytic judgement that does not also express a fact of essence is incomprehensible. And so, even if the first two moves of the error theory partly explain away our intuitions, this third move does not, and so the problematic asymmetry remains.

If some such case could be presented, this objection would be met. I think such cases can be presented, though, of course, they will be controversial. Here's one case: 'Jaywalking is a crime.' This statement is analytic (in Kant's sense) in that being a crime is contained in our concept of jaywalking. Nonetheless, it is not essential to jaywalking that it is a crime. This is a contingent fact; it is only a crime given that it endangers pedestrians and motor vehicilists. But in a society in which this was not the case, jaywalking need not be a crime at all. Another case: 'Mary's biological child was born of her mother, Mary.' Plausibly, our concept of biological child (covertly) contains within it the concept of being begat by that child's mother. Nonetheless, with the rise of surrogate pregnancy, we can see that this is by no means essential to Mary's biological child. Mary's biological child could just as well have been born by someone else, say Mary's cousin Jill. And so, here is another case of a proposition that is analytic (in one sense) and yet whose predicate does not express a property that essentially holds of the object expressed by its subject.[7]

By appealing to these cases, the sparse modalist can apparently capture our clear intuition that there's something asymmetric between the way

[7] Despite the fun Kripke made of it, Kant's example that 'Gold [the element] is gold [in color]' is probably another case. Maybe a moral to be drawn here is that sometimes, in light of learning new things about some phenomenon, we broaden the extension of its corresponding concept without always (or as quickly) broadening that concept's comprehension too.

Socrates relates to {Socrates} and the way {Socrates} relates to Socrates; but it's no essential asymmetry. And so, here's our error theory.

8. Skiles's first objection

SMA seems to fare much better than EMA. However, Skiles (2015) has recently argued that, even if SMA succeeds over EMA in this respect, it fails in another. SMA is extensionally inadequate too. But, unlike EMA, it fails in both directions: it's both too broad and too narrow.

Against the necessity of SMA, Skiles presents four (apparent) cases of abundant essences. First, it is intuitively true that the Eiffel Tower is essentially a tower. But, plausibly, being a tower is an abundant property. Second, it is intuitively true that the Eiffel Tower was essentially designed by Gustave Eiffel. But, plausibly, being designed by Gustave Eiffel is abundant. Third, it is intuitively true that the proposition <The Eiffel Tower is a tower> is essentially true if the Eiffel Tower is a tower. But, plausibly, being true if the Eiffel Tower is a tower is abundant. And finally, it is intuitively true that Socrates is essentially self-identical to himself. But, since distinctiveness properties were classified as abundant, so too must identity properties. Therefore, Skiles concludes, whereas EMA was too broad, SMA is far too narrow. It can't make sense of abundant essences.

Skiles's too-narrow objections to sparse modalism are plausible. Nonetheless, each of his purported counterexamples is problematic, for one reason or another. With respect to the first case, the sparse modalist can deny that being a tower is abundant. Being a tower is a property that a science might appeal to in order to describe the world without redundancy; certain causal powers are intuitively associated with being a tower; and the world does appear carved up by towers in the important kind of way. So, plausibly, it's sparse. With respect to the second case, the sparse modalist can deny that the Eiffel Tower is essentially designed by Gustave Eiffel. Even Fine supposed that essential properties put de re modal constraints on their bearers, such that the bearers could not exist without those properties. But, surely, the Eiffel Tower could exist in a world where Gustave Eiffel does not. Every artist would like to think that they and only they could have created their work; but this is not so.

Finally, with respect to all four cases, but especially with respect to the third and fourth, I think the sparse modalist can respond that Skiles has abused our intuitions here. Consider his remarks: "These are plainly not incidental features… [T]hey pertain to the very nature of these things that any adequate philosophical account of what they are must accommodate and illuminate" (Skiles 2015: 11). Skiles supposes that, if the sparse modalist does not count these properties as essential to their bearers, then she must count them as accidental to them, taken in the usual sense such that they could exist without them. But this is not so, as we have seen in a previous section. The sparse modalist can still help herself to properties that are nonessential but nonetheless necessary for the existence of the object under consideration. Utilizing this resource, she can offer another easy error theory again. These properties are not essential to their bearers (because not sparse). Nonetheless, they might reasonably seem essential to them. This is because of the influence of pure modalism. But, Fine has correctly shown us that this is wrongheaded. Those intuitions, nonetheless, remain. And so, she can say that Skiles has made the same mistake as the pure modalist in mistaking mere necessary-for-existence properties with essential ones.

9. Relativized Sparse Modalism

All of these error theories might seem less than perfectly satisfying. Suppose one's intuitions persists that {Socrates} essentially has Socrates as a member, and that the Eiffel Tower is essentially a tower. Perhaps these intuitions really need to be explained, not explained away, as I have done.

De Melo (2019) has recently offered one way the modalist might revise her account again to capture these intuitions. De Melo's basic move is to suppose that whether or not a given property or relation is sparse cannot be determined absolutely. Rather, sparsity comes in degrees, and whether or not any given property (intrinsic or extrinsic) is sparse is always determined (in part) by what kind of thing the object is to which we are considering ascribing it. In this way, de Melo distinguishes between aSMA – the absolute Sparse Modal Account, what we've just been calling 'SMA' – and rSMA – the relativized Sparse Modal account. He recommends the sparse modalist go in for rSMA over aSMA.

[a]SMA, as Wildman (2013) details it, is comprised of two theses: sparse property modalism ([a]SPM), and sparse relation modalism ([a]SRM). Like Wildman, de Melo wants to allow that some relations can count as sparse. And so, he recommends the sparse modalist adopt alternative dual theses, rSPM and rSRM:

> [rSPM] Object x essentially has property Φ iff (1) necessarily, x has Φ if it exists; and (2) Φ is a sparse property relative to a kind, K, that includes x.
>
> [rSRM] Object x essentially bears relation Ψ to y iff (1) necessarily, if x exists, then Ψ holds of the ordered pair <x, y>; and (2) Ψ is a sparse relation relative to the 'slot' of the ordered pair occupied by x and relative to a kind, K, that includes x.

With these revisions to sparse modalism, de Melo says, no error theories are needed. The sparse modalist can better respond to both Fine and Skiles.

Reconsider Fine's {Socrates} and Socrates case. As noted, there's supposed to be a kind of essential asymmetry here, which aSMA was unable to capture. But rSMA, it seems, can capture that intuition. We can say that {Socrates} essentially has Socrates as a member because membership relations are sparse relative to the kind of thing {Socrates} is – namely, a set. And so, that checks out. And, moreover, we can also say that Socrates is not essentially a member of {Socrates}. This is because membership relations are not sparse relative to the kind of thing Socrates is – namely, a human being. Properties like being rational and relations like originating from (a certain) zygote, ζ, are plausibly sparse relative to the humankind; but set membership relations are not. That checks out too.

Now reconsider Skiles's case of the Tower. I have said that being a tower (maybe) is sparse per our liberal criteria stated earlier. But, then again, maybe it isn't. rSMA has the virtue of not needing to worry about whether or not it is sparse absolutely. In this case, rSMA can appeal to the kind of thing the Eiffel tower is. Plausibly, the kind of thing that the Eiffel tower is an artifact. And relative to artifacts, being a tower is sparse. It is a property that artifacts naturally might possess; moreover, being a tower carves up the world of artifacts at its joints. And so, even if being a tower would not be a natural property relative to other things, relative to the Eiffel tower, it is. And so, rSMA can capture our intuition that the Eiffel

Tower is essentially a tower too. What has been said here goes also for Skiles's other cases, such as the proposition and its truth conditions. Relative to the kind, proposition, having truth conditions is perfectly sparse – it carves up the world of propositions at its joints and grounds qualitative similarity between different propositions.

10. Skiles's second objection

Bracket rSMA for now. More promisingly, Skiles has argued that [a]SMA fails in the other direction too. It is extensionally inadequate because, like EMA, it is still too broad. I think he's right on that point, but I think his particular counterexamples do not work. He offers two. First, it is intuitively true that, for any given water molecule, μ, μ essentially has an oxygen atom as a part. Having an oxygen atom as a part seems to count as a sparse relation per SRM. But, if this is so, then it will also be essential to μ that it have both an oxygen atom as a part and is such that $2 + 2 = 4$. But, being such as to have an oxygen atom as a part and be such that $2 + 2 = 4$ is intuitively not essential to μ. And so, SMA is too broad.

Immediately after presenting this case, Skiles considers a potential sparse modalist response. He imagines that the sparse modalist will deny that this case counts as a genuine counterexample because, even while the simple property of having an oxygen atom as a part is sparse, the conjunctive property of being such as to have an oxygen atom as a part and be such that $2 + 2 = 4$ is not sparse. She can say that, if complex properties or relations are admitted, then they must be restricted to complexes only containing other sparse properties or relations. Being such that $2 + 2 = 4$, we have said, is not sparse.

So, Skiles offers another counterexample. Second, he says, consider the disjunctive property of being such as to have an oxygen atom as a part or being human. We have already said that both disjuncts of this disjunctive property are sparse. And so, it follows that this disjunctive property is sparse too. But, intuitively, it is not essential to μ that it have an oxygen atom as a part or is human. Paraphrasing Fine, there is nothing in the essential nature of μ that connects it to humanity.

Intuitively, the sparse modalist should not want to rule out all disjunctive or conjunctive properties. For example, it does seem essential to water molecule μ that it have the conjunctive property of being such that it has an oxygen atom as a part and has two hydrogen atoms as parts. And so, prima facie, Skiles's second case, at least, is successful.

However, the main problem with Skiles's second critique is that it misrepresents the sparse modalist's response to his first (too broad) counterexample. In response to his first case, the sparse modalist says by way of clarification that a complex property is sparse only if all of the properties or relations of the complex are also sparse. From this, Skiles's then infers that the sparse modalist must then count the disjunctive property of being such as to have an oxygen atom or be human as sparse. But the sparse modalist need not count this. The sparse modalist says a complex property is sparse only if all of its parts are sparse; she does not say that a complex property is sparse if all of its parts are sparse. The sparse modalist can allow that some complex properties are sparse, while also saying that each complex property must be taken on a case by case basis. In fact, this is what she should say, given the (scientific) criteria of sparsity.

I think, therefore, that both of Skiles's objections to sparse modalism fail. Nonetheless, as indicated, I am sympathetic with his second critique. Like him, I think that sparse modalism is still too broad, though not on account of the cases he offers. A better counterexample is available here.

11. Sparse modal propria

To introduce the new counterexample, it would be helpful to first begin by considering a classical distinction between types of properties that an object might possess.[8] Historically, philosophers made a tripartite distinction between an object's essential properties, its accidental (or coincidental)

[8] Alternatively, this might be put in terms of "ways of having a property" (Leslie 2011: 277). There is an ongoing essentialist dispute about whether the essence vs. accident distinction describes kinds of properties or ways of property instantiation.

properties, and importantly, its propria.[9] Like Fine, metaphysicians of antiquity understood an object's essence or essential properties to be those properties that make it to be what it is (*id quo res est id quod est*). They are the most metaphysically significant properties that the object possesses, which it must retain throughout all change lest it undergo substantial corruption. By contrast, they understood an object's accidental properties to be those properties that do not make it to be what it is, and which an object either does or might change throughout the duration its concrete existence. And finally, the medievals understood a proprium to be a property that an object possesses throughout the entirety of its concrete existence, one that, in fact, it could not come to exist without, but which nonetheless does not make it to be what it is. A classic example of a proprium of Socrates is his property of being capable of getting a joke (i.e. his risibility). Other (potential) examples of an object and one of its propria include: a triangle and its having interior angles summing to 180 degrees; a duck and its having webbed feet; salt and its solubility in water; and so on.

There are two points I want to make in drawing this classic distinction. The first is that we should recognize this tripartite division among properties, and that we should want our best account of the character of essential properties to recognize it too. This point is intuitive from the first-person perspective; even if Socrates goes his whole life being two-footed, we should not want to say that it was therefore essential to him, since he never existed without it. Additionally, it is intuitive from the perspective of the philosophical tradition; we should want to preserve the distinction made by Aristotle, the medievals, and other metaphysicians as well.

And the second point is critical. The medievals cast the distinction between essences, propria, and accidents in (partially) temporal terms. But there are also modal counterparts of propria.[10] And importantly, some

[9] Cf. Porphyry of Tyre's (c. 270 AD) Isagoge (esp. Chapter 4, "Of Property"). For a recent extended treatment of this concept, see Bassford (2021).

[10] Fine (1995) has recognized this distinction too. His way of carving out the distinction between essences and propria, however, differs in certain respects from the way I'll be drawing it. He considers propria a kind of consequential essence (as opposed to a kind of constitutive essence), whereas I deny that they are essences at all. See Chi (2020) for more on this point.

modal propria are sparse. For this reason, there are a host of counterexamples to sparse modalism, since it cannot distinguish propria of this variety from essences.

Here's a counterexample. Necessarily, Socrates has the disjunctive property of being blind or sighted if he exists. In all possible worlds, he is either sighted (he sees) or he is blind (sight has been deprived of him). Moreover, being sighted or blind is sparse; it carves up nature at its joints since (a) it grounds genuine qualitative similarity and dissimilarity – namely the qualitative resemblance between visual creatures; (b) it tracks causal powers – namely, the power of sight, for to have it just is to have the natural capacity to see; and (c) it is utilized in our best sciences – namely, it is used to define the discipline of optometry, the scientific study of creatures that have this property, and how best to maintain the proper functioning of the former disjunct while preventing the emergence of the latter (via the eyes). Nonetheless, Socrates's property of being sighted or blind is not essential to him, as such. And so, for this reason, absolute sparse modalism (aSMA) is too broad. It counts as essential properties properties that are only sparse modal propria of their bearers.

Relativized sparse modalism (rSMA) falls to this counterexample as well. The relativized sparse modalist might have been inclined to respond to Skiles's too broad counterexample of water molecule μ being such as to have an oxygen atom as a part or be human, by objecting that being human is not sparse relative to the kind of thing a water molecule is. And so, she might have resisted this counterexample by supposing that it fails the sparse modalist's stipulation that a complex property is sparse only if all of its conjuncts or disjuncts are sparse relative to the kind to which the object belongs. But this response will not help her with respect to this sparse modal propria counterexample. Using de Melo's language, being sighted or blind is a property that it is perfectly natural for the kind of thing that Socrates is to possess. In this way, it will meet both criterion (1) and (2) of rSPM too. But, the same point stands. Socrates is not essentially sighted or blind. And so, for this reason, relativized sparse modalism (rSMA) is too broad as well. It counts as essential properties properties that are only sparse modal

propria of their bearers.[11] In this way, the case presented is problematic for the sparse modalist, whether she supposes that sparsity is absolute or relative to the object's kind.

12. A potential response

I think that this counterexample points to an endemic problem with all varieties of modalism up to the present point, one which demands serious revision. I'll offer my diagnosis to that effect momentarily. But before doing so, it's worth first considering a potential response to my critique.

It might be objected that my purported counterexample does not show that sparse modalism is extensionally inadequate. Rather, what it shows is that we ought to exclude disjunctive properties from counting as candidate essential ones, even if being sighted or blind seems like it might be sparse. The response to Skiles's critique has shown us that, perhaps, we ought to count certain conjunctive properties as sparse, such as being such as to have an oxygen atom as a part and such as to have two hydrogen atoms as parts. This is okay because it is essential to water molecule μ that it have this property. But no example of a disjunctive property that is essential can likewise be offered. And so, SMA may be saved by simply rejecting disjunctive properties as counting as essential properties wholesale.

Prima facie, this is a smart defense on behalf of SMA. However, I think this response is unsatisfactory because too rash. If the sparse modalist would have us exclude all disjunctive properties as candidate essential ones, then she would also have us exclude all determinable properties from counting as candidate essential ones too. This is because determinable properties are both extensionally and intensionally equivalent to the disjunctive property comprised of all of its determinates. Consider the property being red. On its face, this doesn't look like a disjunctive property. But, there's good reason to think it really is disjunctive. Consider that no object is red simplicter.

[11] In certain respects, in relativizing sparsity, rSMA fares even worse that aSMA here. Consider Socrates's (relative) sparse proprium of being such that, for any given joke, he either gets it or misses it (which is not equivalent to simply failing to get it – a boulder might fail to get it a joke, but it never misses one).

An object has the property of being red iff it has the property of being maroon or auburn or mahogany or burgundy or coquelicot… etc.

But, we would want to count certain determinables as being essential to their bearers, as such. As a pointed example, consider the property of being an animal. Like being red, nothing has the property of being an animal simplicter. To be an animal is to be a mammal or a reptile or a fish… etc. But Socrates is essentially an animal. And so, if we exclude all disjunctive properties as candidate essential ones, we must exclude all determinable properties too. If we exclude all determinable properties as candidate essential ones, we must exclude the property of being an animal. But certain objects do have the property of being an animal essentially. And so, we ought not exclude all disjunctive properties from being candidate essential ones to their bearers. The objection stands.

13. Concluding remarks

I conclude, therefore, that sparse modalism (SMA) is an extensionally inadequate analysis of the character of essential properties. It is a genuine improvement over EMA, but like pure modalism, it is still too permissive in what it counts as candidate essential properties. The best option would be for the modalist to once again revise her account in such a way as to circumvent the counterexample of sparse modal propria. How she ought to revise her account at this point in the dialectic is unclear. Strategies for doing so will depend on the proper diagnosis of what has gone wrong. Before concluding, I'll offer my own diagnosis.

I think, ultimately, SMA has gone wrong because it offers no explanation of how all of the properties that an object possesses are intimately related to one another. We have said that the essential properties of an object are supposed to come with a kind of metaphysical gravitas. By adding in a sparse criterion, SMA was able to partly satisfy that intuition, for now the essence of the thing will track its causal powers, inform us about how it relates to other kinds of things, etc.

However, knowing the essence of a thing is supposed to provide us with at least one other kind of explanation about that thing's characteristics. Namely, the essence of a thing should make sense of why the thing has

all properties that it does, whether potential or actual. This intuition, as indicated, was shared by Aristotle (c. 350 BC[c], Post. An. 74b5ff) and the medievals. However, it was also shared by many of the early moderns too. Consider Locke's (1689) point that the "very essentia, or being, of the thing itself [is] that foundation from which all its properties flow, and to which they are all inseparably annexted" (III.iii.18). Consider also Neoscholastic philosopher C. N. Bittle (1939, 117), writing before the advent of modalism, who remarked that "Out of the essence as out of a matrix all being of a thing is, so to say, born… [T]he elements of a thing, which constitute its being, have existence only in so far as they flow (are born) from the essence." Finally, consider also Kelley's remarks that an object's

> essential attribute causes or explains the existence of [its] other attributes… [T]he 'lub-dub' sound is a superficial trait; it is merely a by-product of the heart's essential function, which is to circulate the blood. This essential function explains many of the heart's other properties: the way it beats, the way it is hooked up to the veins and arteries, even the sound it makes. But explanation is a one-way street. The 'lub-dub' sound does not explain the heart's function. (Kelley 1998, 40)

Essences (at least partly) explain all of a thing's properties. At present, SMA does not capture this general intuition. It leaves why the object has all of the essences that it does brute; it leaves why it has the propria that it does brute; and it leaves why it has the accidents that it does brute, as well. Here, I think, is the account's endemic weakness.

Modalists have worked recently to make their accounts mimic in certain respects Fine's RDA. But they ought not forget that there's at least one other account of essences that needs reckoned with too, namely Gorman's (2005) Ontic Explanation Account (OEA). On such an account, H is essential to s iff s's possessing H (ontically) explains s's possessing the other properties that it does too. Accounting for propria is (presumably) no problem there. And so, my advice for the modalist is that she would do well to mimic some features of his account, as well. Maybe a modalism of some such form can be constructed. Maybe 'explanatorily basic,' once added as a criterion into SMA, could even be cashed out modally, in terms of strict

implication relations or iterated counterfactual relations. The upshot of this paper is that more works remains to be done.

Acknowledgements

For useful comments on this paper, and discussion of topics surrounding it, the author owes a debt of gratitude to Ron Avni, Ray Buchanan, Sinan Dogramaci, Dan Dolson, Jon Litland, and Ross Preuss Greene. Thanks!

References

Aristotle. c. 350 BC[a]. "Metaphysics" in *The Basic Works of Aristotle*. W. D. Ross (trans); R. McKeon (ed). Random House, 1941.

Aristotle. c. 350 BC[b]. "Topics" in *The Basic Works of Aristotle*. W. A. Pickard-Cambridge (trans); R. McKeon (ed). Random House, 1941.

Aristotle. c. 350 BC[c]. "Posterior Analytics" in *The Basic Works of Aristotle*. G. R. G. Mure (trans); R. McKeon (ed). Random House, 1941.

Armstrong, David M. 1979. *A Theory of Universals*. Cambridge UP.

Bassford, Andrew Dennis. 2021. "Essence, Effluence, and Emanation: A Neo-Suarezian Analysis." *Studia Neoaristotelica* 18(2): 139–86. https://doi.org/10.5840/studneoar20211827

Bittle, Celestine N. 1939. *Ontology: The Domain of Being*. Bruce Publishing.

Chi, Zeyu. 2020. "Essence, Propria, and Essentialist Explanation." *History of Philosophy and Logical Analysis* 23(2): 338–61. https://doi.org/10.30965/26664275-02302004

Cowling, Sam. 2013. "The Modal View of Essence." *Canadian Journal of Philosophy* 43(2): 248–66. http://dx.doi.org/10.1080/00455091.2013.827395

De Melo, Thiago Xavier. 2019. "Essence and Naturalness." *The Philosophical Quarterly* 69(276): 534–54. https://doi.org/10.1093/pq/pqz014

Descartes, René. 1647. "Comments on a Certain Broadsheet" in *The Philosophical Writings of Descartes* (Volume 1). J. Cottingham; R. Stoothoff; & D. Murdoch (trans / eds). Cambridge UP, 1996.

Dunn, J. Michael. 1990. "Relevant Predication 3: Essential Properties." in *Truth or Consequences*: 77–95. J. M. Dunn & A. Gupta (eds). Kluwer Academic Publishers.

Fine, Kit. 1994. "Essence and Modality." *Philosophical Perspectives* 8: 1–16. https://doi.org/10.2307/2214160%E2%80%A2

Fine, Kit. 1995. "Senses of Essence" in *Modality, Morality, and Belief: Essays in Honor of Ruth Barcan Marcus*: 53–73. W. Sinnott-Armstrong (ed). Cambridge UP.

Gorman, Michael. 2005. "The Essential and the Accidental." *Ratio* 18: 276-289. https://doi.org/10.1111/j.1467-9329.2005.00290.x

Kant, Immanuel. 1781. *Critique of Pure Reason*. N. K. Smith (trans / ed). MacMillan & Co., 1958.

Kelley, David. 1998. *The Art of Reasoning*. W. W. Norton & Co.

Kripke, Saul A. 1972/1980. *Naming and Necessity*. Harvard UP.

Leslie, Sarah-Jane. 2011. "Essence, Plenitude, and Paradox." *Philosophical Perspectives* 25: 277–96. https://doi.org/10.1111/j.1520-8583.2011.00216.x

Lewis, David K. 1983a. "New Work for a Theory of Universals." *Australasian Journal of Philosophy* 61(4): 343–77. https://doi.org/10.1080/00048408312341131

Lewis, David K. 1983b. "Preface" to his *Philosophical Papers* (Volume 1). Oxford UP.

Livingstone-Banks, Jonathan. 2017. "In Defense of Modal Essentialism." *Inquiry* 60(8): 816–38. https://doi.org/10.1080/0020174X.2016.1276855

Locke, John. 1689. *An Essay Concerning Human Understanding*. R. Woolhouse (ed). Penguin, 2004.

Plantinga, Alvin. 1970. "World and Essence." *The Philosophical Review* 79(4): 461–92. https://doi.org/10.2307/2184289

Plantinga, Alvin. 1979. "De Essentia." *Grazer Philosophische Studien* 7: 101-121. https://doi.org/10.5840/gps19797/819

Porphyry of Tyre. c. 270 AD. Isagoge. O. F. Owen (trans); R. Pearse (ed). Christian Classics Ethereal Library.

Robertson, Teresa; & Atkins, Philip. 2018. "Essential vs. Accidental Properties" in *The Stanford Encyclopedia of Philosophy*. E. N. Zalta (ed). Available online at plato.stanford.edu

Roca-Royes, Sonia. 2011. "Essential Properties and Individual Essences." *Philosophy Compass* 6(1): 65–77. https://doi.org/10.1111/j.1747-9991.2010.00364.x

Schafer, Jonathan. 2004. "Two Conceptions of Sparse Properties." *Pacific Philosophical Quarterly* 85: 92–102. https://doi.org/10.1111/j.1468-0114.2004.00189.x

Skiles, Alexander. 2015. "Essence in Abundance." *Canadian Journal of Philosophy* 45(1): 100–12. https://doi.org/10.1080/00455091.2015.1013173

Stalnaker, Robert. 1979. "Anti-Essentialism." *Midwest Studies in Philosophy* 4(1): 343–55. https://doi.org/10.1111/j.1475-4975.1979.tb00385.x

Vetter, Barbara. 2009. "Essence and Potentiality." Presented at the Graduate Center of the City University of New York. Available online at academia.edu.

Wildman, Nathan. 2013. "Modality, Sparsity, and Essence." *The Philosophical Quarterly* 63(253): 760–82. https://doi.org/10.1111/1467-9213.12059

Wildman, Nathan. 2016. "How (Not) to be a Modalist about Essence" in *Reality Making*: 177–96. M. Jago (ed). Oxford UP.