The Newman Problem of Consciousness Science

Johannes Kleiner[1,2,3,4]

[1]Munich Center for Mathematical Philosophy, LMU Munich
[2]Graduate School of Systemic Neurosciences, LMU Munich
[3]Institute for Psychology, University of Bamberg
[4]Association for Mathematical Consciousness Science

Abstract. The Newman problem is a fundamental problem that threatens to undermine structural assumptions and structural theories throughout philosophy and science. Here, we consider the problem in the context of consciousness science. We introduce and discuss the problem, and explain why it is detrimental not only to structuralist assumptions, but also to theories of consciousness, if left unconsidered. However, we show that if phenomenal spaces, and mathematical structures of conscious experience more generally, are understood in the right way, the Newman problem does not arise. The upshot of our paper is that consciousness science needs to be careful in which definition of structure to consider, but if it is, the Newman problem disappears.

1. Introduction

The Newman problem is a fundamental problem for structural theories and structural assumptions throughout science. It was first raised by Newman (1928) in response to Russel's *The Analysis of Matter*, and concerns theories or assumptions which posit that:

> " 'There is a relation R such that the structure of the external world with reference to R is W.' " (Newman, 1928)[1]

Here, R denotes what would now be called the type of a structure. This could, to take a very simple example, be a partial order relation. W is a specification of such structure, meaning that it provides a set of mathematical objects—the elements that are to be related—and specifies which elements in that set are related by the binary relation.

The problem with such postulate is that "[a]ny collection of things can be organised so as to have the structure W, provided there are the right number of them. Hence the doctrine that only structure is known involves the doctrine that nothing can be known that is not logically deducible from the mere fact of existence, except ('theoretically') the number of constituting objects" (ibid.).

It should not be immediately clear or self-evident why the antecedent of this statement is true—why any collection of things can be organised so as to have any structure W, provided there are the right number of them—, and neither why, if the antecedent is indeed true, it constitutes a problem. We explain why it is true, and why it does constitute a problem in Section 2.

Consciousness Science—the scientific investigation of conscious experience and its relation to the physical domain—is currently seeing early signs of a structural turn (Kleiner, 2024). As we will explain in Section 3, the Newman problem can undermine structural research and hence needs to be addressed for any theory-based structural research program to go ahead as intended.

[1]"The world consists of objects, forming an aggregate whose structure with regard to a certain relation R is known, say W; but of the relation R nothing is known (or nothing need be assumed to be known) but its existence; that is, all we can say is, 'There is a [type of] relation R such that the structure of the external world with reference to R is W.' " (Newman, 1928).

The goal of this paper is to show that phenomenal spaces, and similar explications of the mathematical structure of conscious experience, do *not* suffer from the Newman problem, if the mathematical structure of conscious experience is understood in the right way. Nothing hinges on the particularities of consciousness here, other than that the methodology of structural claims that resolves the Newman problem was introduced in the context of consciousness. Therefore, we hope that this work might be of interest also to those who work on structural questions independently of consciousness.

1.1. **Previous work.** The Newman problem is almost 100 years old. Hence, it is no surprise that there is a large body of literature on the topic that discusses and clarifies the problem, as well as a host of different possible resolutions. We locate the work presented here in the landscape of existing resolutions in Section 6 and recommend (Frigg & Votsis, 2011) for an excellent review thereof.

In consciousness science, too, the problem has been discussed and resolutions have been proposed.

Lyre (2022) addresses the Newman problem in the context of a proposed relation between brain states and experiences called *Neurophenomenal Structuralism*. Here, the Newman problem threatens to undermine the claim that neural structures represent the structures of worldly states and processes. It constitutes a problem about what a subject can know about the world, so to speak. Lyre proposes a solution for the Newman problem that follows Russel's own answer to Newman (Russell, 2014), "that certain spatiotemporal [relations in the domain of worldly states and processes] do indeed carry over to [relations among neural states and processes]. We can indeed directly refer to certain spatiotemporal [relations in the domain of worldly states and processes]—or, in Russell's words, are 'directly acquainted' with them." (Lyre, 2022) That is the case, according to Lyre, because the sense organs encode the very spatiotemporal relations that govern external states, for example spatial changes or temporal differences.

Lyre's proposal targets the ramifications of the Newman problem for individual subjects and their epistemic or representational capacities. This paper, in contrast, is concerned with the abstract case of structural claims as part of scientific or philosophical theorizing.

Chalmers (2022) explains the Newman problem when applied to phenomenal consciousness in the context of his comparison of Carnap's logical construction of the world and Lewis's account of Humean supervenience. Chalmers endorses Carnap's resolution of the problem in terms of naturalness conditions (cf. Section 7.2), and concludes that "[b]ecause of Newman's problem, any construction system needs something extralogical in the base" (Chalmers, 2022). Our result, albeit not spelled out in terms of the systems applied by Carnap or Lewis, challenges this claim.

The Newman problem also surfaces in the discussion of consciousness' potential intrinsic properties. Both (Seager, 2006) and (Brüntrup, 2011), for example, take the Newman problem to show that consciousness must be taken to exhibit intrinsic properties or intrinsic qualities. "It is very satisfying to see that the intrinsic nature argument is exactly what is required to avoid Newman's problem, and one would want it to be the case that both Russell and Eddington's deployment of consciousness as an intrinsic nature was explicitly directed at this issue" (Seager, 2006).

What our paper adds to this research is the proposal that if spaces and structure of conscious experience are understood in the right way from the start, no further resolution of Newman's problem is required.

1.2. **Structure of this paper.** After explaining Newman's problem in Section 2, we discuss it implications for consciousness science in Sections 3 and 4. Section 5 is devoted to explaining how a suitable definition of phenomenal spaces, and of mathematical structure of conscious experience more generally, avoids the Newman problem from the start. Section 6 explains how this generalizes to structural claims that do not target consciousness. Full mathematical details of the resulting general proposal are given in Appendix A. Section 7 discusses the limits of the methodology we introduce, and concluding remarks are offered in Section 8.

2. The Newman Problem

Newman's problem arises because of what expressions like

$$\text{“the structure of the (...) world with reference to } R \text{ is } W\text{”} \quad (2.1)$$

(Newman, 1928) are traditionally taken to mean. For a mathematical structure W that comprises a domain C (the 'elements' of the structure) and a relation R , this traditional meaning consists of the following two conditions:

(D1) The elements of the domain C are properties of the world.

(D2) The relation R as specified by W exist.[2]

Because a mathematical relation R is a collection of tuples of elements of the underlying domain, Condition (D2) is actually stating that the tuples that constitute the relation R as specified by W exist.

Expressions like (2.1) are taken to be true if and only if (D1) and (D2) are true. Other than (D2), "of the relation R nothing is known (or nothing need be assumed to be known) but its existence" (ibid.). That is the content of (2.1) as traditionally conceived.

Let us consider, as an example, a partial order structure. Mathematically speaking, a partial order structure consists of a set of elements C, called the domain of the structure, on the one hand, and a binary relation R, on the other hand. The binary relation R is a subset of $C \times C$, meaning that it is a collection of tuples of the form (c_1, c_2), usually written as $c_1 \leq c_2$ in the case of partial orders. The fact that the partial order structure W consists of these two constituents is often expressed by writing $W = (C, R)$.

Condition (D1) then states that the elements of the domain C of the partial order (as specified by W) are properties of the world. Condition (D2) states that there exists a binary relation (viz. a collection of pairs of elements) that relates the elements as specified by W. The elements need to be arranged in tuples as specified by W for this condition to be true. This is what is means to say that a partial order structure W is a structure of the world, according to the traditional understanding of expressions like (2.1).

The problem with this understanding of (2.1) is that while elements in the domains of the structure are required to have referents in the world (a structure is a structure of the world only if there are properties as specified in the domain of W), this isn't true of the relation. The relation is not required to have a referent in the world. The condition on the relation is only exposed qua condition on the properties. In other words, the relation is not required to correspond to any concretum in the world, Condition (D2) only relates abstract formalism in W to abstract descriptions of the world. As a consequence, any abstract specification of structure in the world will satisfy (D2). This consequence is expressed by the following theorem, presented in (Frigg & Votsis, 2011; Ketland, 2004).

Theorem 1 (**Newman's Theorem**)**.** Let C be a collection of individuals and let W be a structure whose domain has the same cardinality as C. Then there exists a structure W_C whose domain is C and which is isomorphic to W.

That is to say, independently of whether the world actually comprises a relation R as specified by W, if the properties exist, one can simply define a suitable relation to render Condition (D2) true. "[G]iven any structure, if collection C has the same cardinality as that structure, then there is a system of relations definable over the members of C so that C has that structure. (...) [A]ll we have to do in order to define a relation is to put elements in ordered tuples and put these tuples together in sets, which we can always do as long as we have enough elements" (Frigg & Votsis, 2011). Condition (D2) is not itself depending on anything in the world over and above the dependence already established by (D1). This is the cause of the Newman problem.

There are a number of ways to resolve the Newman problem, cf. (Frigg & Votsis, 2011, Sec. 3.4) for an excellent discussion. When the problem is presented as above, the obvious route to a solution of the Newman problem is to ask whether one could not replace (D2) by a

[2] This requirement is usually implicit in the requirement that the axioms of a structure (transitivity, reflexivity, anti-symmetry, for example, in case of a partial order), have to hold.

better condition, such that Theorem 1 ceases to apply. This route is precisely the one we will take in Section 5, but first we will discuss why the Newman problem applies to consciousness science, and which ramifications it has for consciousness science.

3. The Newman Problem in Consciousness Science

Consciousness science is seeing early signs of what could be a structural turn. Virtually every field that is involved in consciousness science has started to employ mathematical spaces and mathematical structures as means to investigate, model, or measure the phenomenon.[3] In doing so, many different methodologies and ideas are applied, known under various different names, including quality spaces (Clark, 1993; Rosenthal, 2015; Lee, 2021), qualia spaces (Stanley, 1999), experience spaces (Kleiner & Hoel, 2021; Kleiner & Tull, 2021), qualia structure (Kawakita, Zeleznikow-Johnston, Takeda, Tsuchiya, & Oizumi, 2023; Tsuchiya, Phillips, & Saigo, 2022), Q-spaces (Chalmers & McQueen, 2022; Lyre, 2022), Q-structure (Lyre, 2022), Φ-structures (Tononi, 2015), perceptual spaces (Zaidi et al., 2013), phenomenal spaces (Fink, Kob, & Lyre, 2021), spaces of subjective experience (Tallon-Baudry, 2022), and spaces of states of conscious experiences (Kleiner, 2020). All of these proposals attribute mathematical structure to conscious experiences, which is why we will use the term 'mathematical structure of conscious experience' as an umbrella term to refer to these and similar proposals.

In all of these proposals, there is a mathematical space or mathematical structure E that is claimed to describe, represent or model conscious experience. Modulo terminological choices, all of these proposals endorse some variant of the claim that

$$\text{“the structure of conscious experience is } E\text{”.} \tag{3.1}$$

In (Kleiner & Ludwig, 2024), we have analyzed those proposals that work with explicit conditions to assert such claims, cf (Kleiner & Ludwig, 2024, Sec. 1). Perhaps unsurprisingly, other than explicit statements of the axioms that a mathematical structure is required to satisfy, these are exactly Conditions (D1) and (D2), with 'properties of the world' replaced by 'properties of conscious experiences' or analogous constructs.

In (Kleiner & Ludwig, 2024), we use the term 'aspect' as a placeholder to denote concepts like qualia, qualities, instantiated phenomenal properties, phenomenal distinctions, or similar, that feature in claims of the form (3.1). For terminological simplicity, in this paper, we will work with the concept of *phenomenal properties*, which are properties of the phenomenal character of an experience, where 'phenomenal character' refers to what it is like for an organism to be that organism in a particular state (Nagel, 1974).[4] However, all we say below applies to other conceptual choices (qualia, qualities, phenomenal distinctions, etc.) as well.

In terms of phenomenal properties, we can formulate the conditions placed on structural claims in consciousness science as follows. Expressions like (3.1), where structure $E = (C, R)$ consists of domain C and relation R, are taken to be true if and only if

(C1) The elements of the domain C are phenomenal properties of conscious experiences.

(C2) The relation R as specified by E exists.[5]

To give an example of this in consciousness science, consider a metric space of color

[3]A list of references of current developments is given in Kleiner (2024).

[4]Philosophers often define an experience to be the instantiation of a phenomenal property by an experiencing subject, so that an experience is an event. The phenomenal character of an experience in this framework is what it is like for the subject to undergo such event. Cf. Nida-Rümelin (2018) for more details on and problems of this way of thinking. Those from a more formal context often tend to take the term 'conscious experience' to refer directly to what it is like. In this paper, we will use the term 'phenomenal character' to denote what it is like, in the hope that this choice is the largest common denominator across fields and backgrounds.

[5]As in the case of general structural claims (Section 2), this requirement is usually implicit in the requirement that the axioms of a structure have to hold.

[6]For details on how quality spaces are constructed in consciousness science, cf. (Kleiner, 2024, Sec. 5).

qualities. The requirement that a specification E of such space is a structure of conscious experience—a quality space, for short[6] — comprises, first, the condition that the points of the metric space are color qualities (being phenomenally presented with red, blue, etc.), and that the real numbers that describe the distances in a metric spaces are experienced degrees of similarity of such color qualities (being phenomenally presented with a degree of similarity). Color qualities are phenomenal properties, hence this requirement is Condition (C1). Second, for any two color qualities, there must be an experienced degree of similarity of those color qualities as described by the metric function.[7] This just means that any triple that consists of two color qualities and the corresponding degree of similarity as specified of the metric function must exist. This is Condition (C2). Only if both (C1) and (C2) are satisfied, is this an instance of a quality space.

Because conditions (C1) and (C2) are exactly analogous to conditions (D1) and (D2), the Newman problem applies to consciousness science in the exact same way as it applies in other domains. Explicitly, the Newman problem (Theorem 1) implies that any claim of the form (3.1) is empty, as far as the structural content is concerned. Nothing over and above the cardinality of the set of phenomenal properties is endorsed in a claim like (3.1). This has far-reaching consequences.

4. Implications for Consciousness Science

The Newman problem has a number of ramifications in consciousness science. On the more obvious side of things are its ramifications for structuralist research programs. Less obvious, maybe, is that the Newman problem also undermines work on theories of consciousness.

4.1. **Structuralist Research.** Structuralist research programs in consciousness science come in one of two flavors. They either target the question of what can be known, scientifically or introspectively, about what it is like—what can be known about phenomenal character, in the terminology applied in this paper. Or they target the question of what phenomenal character actually is—in which sense it exists, so to speak. Following terminology of philosophy of science, we might designate the former as *epistemic phenomenal structural realism* (EPSR), and the latter as *ontic phenomenal structural realism* (OPSR). OPSR says that phenomenal structures are ontologically basic: non-structural features of phenomenal character, such as intrinsic qualities, do not in fact exist; only claims of the form of (3.1) can be true. EPSR is the view that all we can know about phenomenal character is its structure; only claims of the form (3.1) can be known. Cf. (Frigg & Votsis, 2011) for the corresponding distinction in philosophy of science. Therefore, if claims like (3.1) are in fact void over and above implications of cardinality, so are OPSR and EPSR. The Newman problem, if unresolved, undermines these research programs. This is well known, cf. e.g. (Lyre, 2022) or (Chalmers, 2023).

4.2. **Theories of Consciousness.** What is less well known, maybe, is that the Newman problem also undermines theories of consciousness. Specifically, it undermines theories that address phenomenal structure, if those theories are intended to be applicable to non-human organisms or non-human systems more generally.

That is the case because for non-human systems, ostensive definitions of phenomenal structures fail. We cannot use language to pick out the referent of a structural claim like (3.1) in non-structural terms, either because non-human systems have no suitable language, or, in the case of LLMs, because they do not use language in the same way as we do. "The ostensive definition [only] explains the use—the meaning—of the word when the overall role of the word in a language is clear. Thus [only] if I know that someone means to explain a colour-word to me the ostensive definition 'That is

[7]A metric space consists of two domains and one function. The domains are the set of points of the space and the real numbers. The metric function maps any two points to one real number. For simplicity, we have formulated (C2) in terms of relations only. Technically speaking, this condition can also be applied to functions because any function $d : C_1 \times C_2 \to \mathbb{R}$ is a unary relation on $C_1 \times C_2 \times \mathbb{R}$. We do think, however, that it is good to distinguish relations and functions in such contexts, and do so in (Kleiner & Ludwig, 2024).

called »sepia«' will help me to understand the word" (Wittgenstein, 1953). In human cases, we can get around purely structural claims like (3.1) by pointing out which phenomenal structure a phenomenal claim like (3.1) is intended to address. In non-human systems, because of the lack of shared meaning of language, this is not an option. The only thing we can do is to specify the structure abstractly, as in (3.1), which is why the Newman problem applies in full force.

This is particularly evident in one of the mathematized theories of consciousness, Integrated Information Theory (IIT) (Albantakis et al., 2023; Oizumi, Albantakis, & Tononi, 2014). IIT comprises a carefully constructed algorithm that specifies, for any mathematical description of a system in a specific state, a complex mathematical structure called Φ-structure (cf. Kleiner and Tull (2021) for a structural exposition of IIT). The Φ-structure is the output of IIT's algorithm. In terms of the terminology applied here, it specifies the phenomenal character that a system is experiencing when it occupies the respective state. But IIT does not provide a phenomenal interpretation of this structure, it only provides the mathematics. This is a perfect example of (3.1).

To provide a Φ-structure is a substantial achievement of IIT. But if the Newman problem applies (which it must if (3.1) is understood as (C1) and (C2)), then IIT's structural claim is entirely void, over and above the cardinally of the elements in the structure.

This is related to what Chalmers (2023) has called the Rosetta Stone problem of IIT: the problem of how to translate the mathematical structure that IIT proposes into phenomenological terms. If Newman applies, it follows that no such translation is possible, as the structural claim is void; IIT's structural claim can always be satisfied simply by defining the required structure over phenomenal properties.

The same applies to other theories of consciousness if they make structural phenomenal claims. Theories are prone to Newman's problem because they are supposed to stand on their own, they should be meaningful independently of ostensive human-language pointers that specify which structure is what in phenomenal character. It should suffice for a theory to specify the phenomenal structure of a system in terms of structural language; the relevant parts of phenomenal character should then be determined.

So how many theories address phenomenal structure and are intended to be applicable to non-human organisms or non-human systems? At present, only a small fraction of theories address phenomenal structure. Examples are IIT, mentioned above, as well as Expected Float Entropy Theory (Mason, 2021) and Rosenthal's quality-space version of higher order thought theory (Rosenthal, 2010). However, it can be argued that addressing phenomenal structure is inevitable once theories start addressing phenomenal character more faithfully than they presently do. Binary distinctions between whether a stimulus is being consciously perceived, or not, or whether a system is conscious at all, or not, might not suffice to explain phenomenal character faithfully (Kleiner, 2024). Furthermore, it can be argued that all theories of consciousness should be formulated in such a way that they can, in principle, be applied to non-human systems or organisms (Kanai & Fujisawa, 2023). This might be part of the desiderata for a theory to count as a meaningful theory of consciousness.

Therefore, the class of theories that will eventually come into the realm of the Newman problem is large. It looks like the Newman problem went by largely unnoticed, as far as scientific theories of consciousness are concerned, because most theories are not advanced enough at the present stage for them to introduce the tools that the Newman problem vexes. But once they do, the Newman problem might well undermine much of the effort in constructing them, if unresolved.

5. Solving the Newman Problem of Consciousness Science

The obvious solution to Newman's problem, when presented as in Section 2, is to replace Condition (D2) resp. (C2) by another condition, so as to modify the meaning that expressions like (2.1) or (3.1) should have, in such a way that Newman's problem ceases to apply. This amounts to proposing alternative *definitions* of expressions like (2.1) or (3.1) that avoid Newman's problem; just like one proposes improved definitions of concepts like qualia or phenomenal consciousness in philosophy of mind to avoid problems the terms might otherwise face.

When improving (D2) resp. ((C2)), a new condition must remain compatible with the spirit of (2.1) resp. (3.1). The major constraints this raises is that the condition should also be formulated abstractly, and must only make use of exist quantifiers ('there exists ...'); no direct reference of properties of the world resp. phenomenal properties can be included. We now discuss three proposals of how this could be achieved. We employ the terminology of phenomenal properties, but the same points could be made with respect to properties of the world, as we explain in Section 6.

5.1. **Higher-Order Phenomenal Properties.** An immediate idea to improve Condition (C2) is to work with higher-order phenomenal properties. Phenomenal properties are properties of the phenomenal character of a conscious experience, and much like first-order phenomenal properties (presumably, for example, being phenomenally presented with red), there are higher-order properties (for example, being phenomenally presented with similarity of two shades of red).[8]

To improve (C2), one could simply add the condition that there exists a higher-order phenomenal property for every relation R in a structure E. That would amount to replaying (C2) by:

(C2′) The relation R as specified by E exists, and there is a higher-order phenomenal property.

The idea is that for every relation R, there is one higher order phenomenal property, and that no two relations can have the higher-order phenomenal property in common.

This condition would indeed resolve Newman's problem because the simple existence of a structure with phenomenal properties as its domain is not sufficient any more to satisfy (C2′). Rather, there must be a phenomenal property (or something in the world, in Newman's terms). This is an additional requirement whose satisfaction does not follow from Theorem 1.

However, (C2′) is not a suitable proposal because the structural phenomenal property that needs to exist has nothing to do with the relation R as specified by the mathematical structure E. The condition does not pin down the relation in any significant sense, over and above the requirement that the number of relations that exist is smaller than the number of higher-order properties. It's not enough to just require *some* phenomenal property to exist.

5.2. **Arity.** In order to remedy the problem of (C2′) that the mathematical structure of E has nothing to do with the higher-order phenomenal property that is required to exist, we have to expand the requirements placed on the higher-order phenomenal property.

While higher-order phenomenal properties do not have, or cannot be taken to have in this context, a mathematical structure that one can simply reference, they do exhibit a feature that in mathematics is called *arity*, and in philosophy may also be called *adicity*. It is the number of lower-order properties the higher-order property is instantiated relative to. For example, if the higher-order property is being phenomenally presented with similarity of two different shades of red, it has an arity of 2.

Relations in the mathematical sense of the term also have arity. It is the number of "slots" in the relation, or in other words, the number of elements that every tuple in the relation comprises. A binary relation, for example, has arity 2 because its tuples are pairs of elements. A relation of arity n comprises n-tuples, each of which consists of a list of n elements of the domain. Making use of this fact, we could modify (C2′) to read:

[8] In (Kleiner & Ludwig, 2024), we have called these 'structural properties', but this choice of terminology might not be ideal as it suggests that these properties already have some structure in the mathematical sense of the term. This is not the case. Rather, they only have *arity* (the number of lower-order properties they are instantiated relative to, cf. Section 5.2 below).

(C2$''$) The relation R as specified by E exists, and there is a higher-order phenomenal property that has the same arity as R.

This is an improvement over (C2$'$) because now the phenomenal property cannot be arbitrary any more.

However, (C2$''$) still fails because there are vastly different relations of the same arity. Arity characterizes a relation to some extent, but it still leaves most details of a relation unspecified.

What is needed to arrive at a satisfying condition is some way of characterizing a relation's mathematical form, that can also be interpreted in terms of phenomenal character.

5.3. **Automorphisms.** One way to characterize a mathematical structure in full (up to a certain point, cf. Section 7) is given by its *automorphism group*. Automorphisms are functions that map every element from a domain of a structure to another element of the domain. The mapping has to be one-to-one (implying that it has to be invertible), and has to preserve the relations (and functions) defined over a structure. If a structure consists of one domain C and one binary relation R, for example, this mapping takes the form

$$f : C \to C \,,$$

and the requirement that it preserves the relation is formaly stated as

$$R(c_1, c_2) = R\left(f(c_1), f(c_2)\right) \qquad (5.1)$$

for all $c_1, c_2 \in C$.[9] Automorphisms form a group because they are invertible, and because any two automorphisms can be concatenated to give a new automorphism.

Automorphisms are intriguing objects in the current context because, once a domain is specified (qua Condition (C1)), a set of automorphism can be specified as a set of functions $\{f_1 : C \to C,\ f_2 : C \to C,\ ....\}$. Neither the relation R, nor the tuples that constitute the relations, have to be specified when specifying the functions in the set.

Of course, if one would only specify a set of automorphisms, the Newman problem would apply just as well. They are formal objects and hence always exist, if the domain contains enough elements. What is needed, in addition, is a link between automorphisms and phenomenal properties. Such a link can be provided, as we now explain.

Let us consider an arbitrary function (also called 'mapping') $f : C \to C$, where C is a domain of a structure that satisfies (C1). An arbitrary mapping can or cannot be an automorphism of a structure E, depending on whether it satisfies the definition of an automorphism, or not—that is to say, depending on what the structure E is, and depending on how elements are mapped by the function. If a function is an automorphism, one often says that it "preserves" the structure. If it is not an automorphism, one says that it does "not preserve" the structure. Those are abstract statements in the domain of mathematics. (Cf. Definition 4 in Appendix A for formal details.)

But in cases where a domain C satisfies (C1), functions $f : C \to C$ can also be understood as something concrete: they describe how phenomenal properties change. To give a very simple example: if a subject has an experience of seeing red, and that changes to an experience of seeing blue, this can be described as a (partial) function that maps from phenomenal properties to phenomenal properties; it maps being phenomenally presented with red to being phenomenally presented with blue. Such a variation of phenomenal properties must, in turn, be understood as a variation of the underlying experience whose phenomenal properties are at issue. Variations of experiences are changes from one experience to another, and for every such change, there is a corresponding variation of (instantiated) phenomenal properties.

Because functions can be interpreted in both abstract and concrete terms, they provide the link between the abstract domain of mathematics and the concrete domain of conscious experiences that is needed to amend Condition (C2$''$). They allow us to express the requirement that the higher-order phenomenal property in (C2$'$) mirror the structure E in terms of behavior of variations as follows: a higher-order phenomenal property must behave as the structure does under variations. This means that the higher-order phenomenal

[9]We write an equal sign here for notational simplicity. The formally correct statement would be $R(c_1, c_2) \Leftrightarrow R\left(f(c_1), f(c_2)\right)$ for all $c_1, c_2 \in C$.

property must prevail in phenomenal character if the mapping between first-order properties induced by a variation preserves the structure; and it must disappear if the mapping between the first-order properties induced by a variation does not preserve the structure. We will say that in former case, the variation "preserves" the higher-order phenomenal property, whereas in the latter case it "does not preserve" the higher-order phenomenal property (cf. Definition 3 in Appendix A for formal details).

We call a higher-order property that satisfies this requirement for a relation R a phenomenal R-property. In concise terms:

(SP) A phenomenal property p is an R-property iff any variation that preserves the relation R preserves p.

We note that this definition is only meaningful if (C1) holds; (C1) provides the "baseline correspondence" between mathematical structure and phenomenal properties that allows to make sense of variations both in terms of changes of phenomenal properties and automorphisms.

We can thus present a suitable extension of (C2) as:

(C2$'''$) The relation R as specified by E exists, and there is a corresponding higher-order phenomenal R-property.

This is the meaning/definition of structural claims like (3.1) we have arrived at, for independent reasons, in (Kleiner & Ludwig, 2024) as well. It makes use of variations which form an important part of earlier proposals of how to define spaces of conscious experiences, for example Rosenthal (2015), and it retains the original Condition (C2): since (C2$'''$) implies (C2), Condition (C2) is a necessary part of Condition (C2$'''$).

Condition (C2$'''$) resolves the Newman problem because the mere existence of some structure is not sufficient to satisfy the condition. The condition requires that there is a phenomenal property of the right sort. This is a requirement whose satisfaction does not follow from Theorem 1. The condition furthermore leaves no freedom for the relation to vary while the property is fixed, as (C2$'$) and (C2$''$) did. Hence it is, as far as we can see now, a viable solution of the Newman problem of consciousness science.

6. A general solution?

In the previous section, we have shown how the Newman problem of consciousness science can be resolved by providing a more careful definition of what structural claims are taken to be. Here, we discuss whether this affords a solution of the Newman problem independently of consciousness.

Before we embark on this discussion, we would like to mention that there are several viable solutions of the Newman problem already, discussed in detail in (Frigg & Votsis, 2011). A review of these solutions would go beyond the scope of this paper, but suffice it to say that the solution presented here might be a case of the 'Real vs. Fictional Relations' class of solutions that attempt to distinguish real relations in the world from those that are merely defined (called 'fictional' by (Newman, 1928)).

The viability of our solution of the Newman problem in consciousness science as a solution of the general Newman problem depends on whether the ingredients we have made use of also exist in the general setting of the Newman problem. These are higher-order phenomenal properties, and the concept of variations.

The notion of higher-order phenomenal property easily carries over to any context in which a structural claim like (2.1) is made. There are higher-order properties of the world in a similar sense as there are higher-order phenomenal properties.[10]

[10] If properties are conceived of as properties of things, one might want to distinguish the concept of higher-order properties from the concept of relational properties. Relational properties are properties between things. There can be both first-order relational properties, and higher-order relational properties. Properties which are properties only of one thing are called monadic properties, and there are both first-order and higher-order monadic properties, the latter of which are properties of properties of one thing. According to this conception of properties, our proposal below could be defined in terms of either relational or higher-order properties, or both; what matters is that the properties in question have arity, also called adicity. I would like to thank Andrew Lee for pointing this out.

The case of variations is more difficult. The crucial property of variations that enables definition (C2$'''$) is that the variation changes both the first-order property, and the higher-order property. Can we make sense of such variations? And if so, what defines the variations that exist as compared to those that do not.

In a context like (2.1), where the reference of a structural claim is "the world", one could make sense of variations in terms of possible world semantics as used in modal logic, cf. e.g. (D. K. Lewis, 1986). Specifically, one could consider the set of all nomologically possible worlds—the set of worlds that are compatible with the laws of nature, that is—and then define a variation simply to be a map from one nomological possible world w_1 (e.g., the actual world) and all its properties to another nomological world w_2 and all its properties. Such a variation preserves a (possibly higher-order) property if and only if it is present before and after the variation, meaning if it is a property of both w_1 and w_2, and it preserves a structure S if and only if it is an automorphism of S, where the latter definition makes use of (D1).[11]

This gives rise to the following exposition of structural claims like (2.1). The notion of R-property is defined as:

(SP) A property p is a R-property iff any variation that preserves the structure R preserves p.

Claim (2.1) is true if an only if the following two conditions are true:

(D1$'$) The elements of the domain C of are properties of the world.

(D2$'$) The relation R as specified by W exists, and there is a higher-order R-property.

While abstract at first, this condition is highly compatible with physical sciences, because nomologically possible worlds should not be understood as atomistic entities. Rather, the set of nomologically possible worlds is intimately connected with initial conditions of natural laws, and a fortiori with repeated experiments. Structural claims so defined can be assessed empirically by considering "chunks" of the actual world in scientific experiments, and by studying how these chunks behave as time or other parameters vary.

An alternative to this approach would be to take properties in the world to be attached to objects, or groups of objects, in the world, and to consider variation of such (groups of) objects. This would also provide a suitable concept because a variation of a (group of) objects would vary both first-order and higher-order properties of the object or group. A definition of this kind would be of advantage because it would be more intuitive as the above. However, it would not naturally align with the foundations of physics, where existence of individual objects (rather than just one global field with particles as modes or excitations thereof) in an intuitive sense is somewhat contested. Still, it might be a viable option, and might actually correspond to the above definition if possible worlds are conceptualized in the appropriate way.

We provide a full formal exposition of our proposal in Appendix A, and discuss a limitation of our approach in Section 7. The consequences of this limitation are, on our view, what ultimately determines the viability of our proposal for purposes of solving the general Newman problem. In the next section, we explain how our proposal relates to the Newman problem when expressed in terms of Ramsey sentences.

6.1. **Ramsey sentence formulation of the Newman problem.** The Newman problem is often stated in terms of Ramsey sentences, introduced by Carnap (D. Lewis, 1970). In a nutshell, for any theory T that contains observational predicates Q_i and non-observational predicates P_i, one can first form a logical conjunction of all of a theory's postulates/axioms/rules to write the theory as a single formal sentence that is usually denoted as (Frigg & Votsis, 2011)

$$T\left(P_1, ..., P_m, Q_1, ..., Q_n\right) . \tag{6.1}$$

The Ramsey sentence of such theory is the result of replacing all non-observational predicates P_i by variables, which we denote as X_i,

[11] Because properties can disappear from w_1, as in the case of consciousness, mappings must be understood as partial functions. Because they need not be surjective, properties can appear in moving to w_2. More details on such mathematical subtleties are given in the appendix, and in (Kleiner & Ludwig, 2024).

and adding an existential quantifier over these variables, denoted by '$\exists$' to the sentence.[12] This gives the theory's Ramsey sentence T_R,

$$\exists X_1 \ldots \exists X_m \; T\,(X_1, \ldots, Q_n) \; . \tag{6.2}$$

A Ramsey sentence encodes a theory's full empirical content. Because the predicates P_i are non-observational predicates, they do not have observational consequences over and above their mere existence and role in the theory T. Therefore, a theory and its Ramsey sentence have the same observational consequences. Furthermore, the Ramsey sentence (6.2) follows logically from (6.1). Cf. (Frigg & Votsis, 2011, Sec. 3.3) for more details.

Making use of Ramsey sentences, the Newman problem can be stated as the following theorem. Here, a model of a theory is *u-cardinality correct* if it has the same cardinality as the unobservable predicates of a theory, and *empirically correct* if its empirical substructure is isomorphic to the empirical substructure of the target domain (Frigg & Votsis, 2011).

Theorem 2 (**Cardinality Theorem**)**.** The Ramsey Sentence of theory T is true if, and only if, T has a model S (i.e. $S \models T$) which is u-cardinality correct and empirically correct.

This theorem establishes that "all we can infer from the truth of [a theory's Ramsey sentence] T_R about the unobservable world is a claim about its cardinality" (ibid.), and that "any claim the [Ramsey sentence] may make about the existence of unobservable relations or their formal properties is automatically true (or 'trivially' true, as the point is often put)" (ibid.).

How does our proposal deal with the Ramsey sentence formulation of the Newman problem?

Our proposal amounts to a redefinition of the truth-condition of structural claims. According to the received view of such truth-conditions, a second-order predicate that expresses a structural claim is true iff Conditions (D1) and (D2) are true. According to our proposal, a second-order predicate that a expresses a structural claim is true iff Conditions (D1′) and (D2′) are true.

This changes the implications of the existential quantifiers in (6.2). They do not assert that there exists structure in the world that satisfies (D1) and (D2), but rather that there exists structure in the world that satisfies (D1′) and (D2′).

As a consequences, the right-to-left direction of Theorem 2 breaks down. While it is still true that the truth of a Ramsey sentence of a theory T implies that there is a model which is u-cardinality correct and empirically correct (Condition (D2) is still a necessary part of Condition (D2′); this is the left-to-right direction of the theorem), the opposite direction fails to hold: it is not the case that any model which is u-cardinality correct and empirically correct implies the truth of the Ramsey sentence, because it also needs to satisfy the R-property condition in (D2′).

As a consequence, with the improved understanding of structural claims that we have proposed above, it ceases to be true that "any claim the RS may make about the existence of unobservable relations or their formal properties is automatically true (or 'trivially' true, as the point is often put)" (ibid.).

7. Objections

In this section, we would like to address one objection to, and one fundamental worry of, our proposal.

7.1. **Reconstructing structure.** The fundamental worry concerns the question of just how much of a mathematical structure can be identified (or "reconstructed") from its automorphism group.

Consider again the three proposals we have made in Section 5. Starting from the fundamental idea to add existential quantifiers of higher-order properties in (C2′), we have subsequently expanded the condition so as to limit the number of mathematical structures that can be associated with a given higher-order property. While Condition (C2′) did not put any constraint on how the structure relates to a higher-order property, Condition (C2″) required arity to match up, and Condition (C2‴) required the variations that constitute a structure's automorphism group

[12] This is an instance of quantification over predicates, which presumes second-order logic.

to match up with the variations that preserve the higher-order property.

The problem we discuss here is that while it is true that an automorphism in general characterizes a mathematical structure in full, it does not do so in extreme situations. Automorphisms "min out" at some point. Once the automorphism group is trivial, it remains trivial even if more structure is added, as we now explain.

Consider a mathematical structure W that consists of a domain C and relation R, where the relation R allows to individuate every element of C uniquely based on relational information alone. This is the case for graphs, for example, if every node of a graph has a unique number of edges that connect to that node, called the degree of that node. The automorphism group of such structure contains only the identity mapping, for every other mapping would not be able to preserve the edge relation (cf. (5.1)). In this case, the automorphism group of the structure is called 'trivial'.

Trivial automorphism groups constitute a problem for our proposal because once the automorphism group is trivial, automorphisms fail to track any further changes to structure that preserve triviality. If, for example, a further edge is added to a graph, while preserving the condition that every node has a unique degree, then the automorphism is trivial before and after the change in structure. It can neither track, nor be used to reconstruct, the difference in structure. Put more abstractly, different relations that are defined over a given (fixed) set of elements can all have the same trivial automorphism group. The condition for there to be an R-property is the same for any relation R that satisfies (D1$'$) for a given set C and whose automorphism group is trivial. This problem of automorphism-based criteria to distinguish structure is well-known in the structural parsimony debate in philosophy of physics (Barrett, Manchak, & Weatherall, 2023).

There are two different responses one can give to this problem, and both apply.

First, one could argue that this problem indicates that Condition (D2$'$) can still be improved. Maybe some more advanced math could be used to resolve structure via automorphisms even if the automorphism group of a structure is trivial. Local automorphisms and sheaves come to mind. Or maybe there is an entirely different way of formulating a condition that replaces (D2). Both are viable options to explore in further research.

Second, one could argue that the problem is not actually detrimental to the proposal, because such relations cannot satisfy both Conditions (D1$'$) and (D2$'$).

To see why this is the case, we first emphasize that the problem we describe here only applies if individual relations already imply that the automorphism group of a structure is trivial. That is because every relation R is required to have a corresponding R-property. If a structure contains more than one relation, and the entirety of them render the automorphism group trivial, all is well.

Consider, therefore, a single relation R that precludes non-trivial automorphisms. For this relation to satisfy (D2$'$), there needs to be an R-property p as described by (SP). This yields two conditions: First, any variation that preserves R needs to preserve p. And second, any variation that does not preserve R must not preserve p. The automorphism group being trivial implies that only the second case applies, so that any variation whatsoever must not preserve p. The mathematical formalism of our proposal implies that if there is to be an R-property p, any world which instantiates the elements in one tuple of R must instantiate p. Because the relation has trivial automorphism, there must at least be two tuples in the relation. Therefore, we must at least have two worlds that instantiate p. But any variation from one of these worlds to the other of these worlds preserves p, as it is instantiated both in the source and target world of the variation. This violates the condition that there is no variation that preserves p, and hence there cannot be an R-property for a relation that has trivial automorphism.[13]

The formal arguments behind this reasoning are provided rigorously in Appendix B (cf. Lemma 2). In summary, the mathematics of our proposal simply deny relations that induce trivial automorphism groups the status of viable objects of a structural claim. This

[13]For a formal proof of this claim, cf. Lemma 2 of Appendix B.

is aligned with the idea that a mathematical structure is only meaningful to the extend that it can be probed by variations.

7.2. **Does naturalness suffice?** One way to resolve the Newman problem is to assume that "only natural relations should be taken into account when pondering the structure of the world; we need not, strictly speaking, deny that the world instantiates (...) any relation compatible with its cardinality, but we submit that only natural relations are taken into account when it comes to assessing the claims of a theory" (Frigg & Votsis, 2011). This idea was introduced by D. Lewis (1983), and is the solution endorsed by Chalmers (2022) in the context of consciousness science, cf. Section 1.1.

Given that our proposal in (D2′) introduces a technical term, one could object that the solution in terms of naturalness is preferable, simply because it is a simpler solution. Is this so?

The solution terms of naturalness amounts to reinterpreting the 'exists' quantifier in (D2). Instead of an abstract existential claim, it would have to be interpreted as quantifying over natural *relations*. If there is a natural relation R as specified by W, then (D2) is true. If not, (D2) is false.

This resolution of the Newman problem is problematic, cf. (Frigg & Votsis, 2011, Sec. 3.4.1(b)). One problem is that what counts as a natural kind might change as science progresses, cf. (Melia & Saatsi, 2006). Another problem is that distinguishing natural kinds from non-natural kinds might require non-structural language of the world to begin with, cf. (Psillos, 2005).

But more fundamentally, even, for this solution to work and be applicable one has to presume that the world is mathematical, and that mathematical terms refer "just like that". One has to presume that it is meaningful to say there is a natural *relation*, where 'relation' is used in the mathematical sense of the term.

This is a substantive claim, and at least when it comes to phenomenal character, there are good reasons to think it is wrong. Conscious experiences do not come with mathematical structure in any meaningful way. Phenomenal character isn't experienced as a metric space, for example. There are experiences, and mathematical formalism is useful to describe or represent experiences. To say that conscious experiences have a mathematical structure is a way of describing them, not part of what they are naturally given as. Similarly, natural kinds (or related concepts) might not constitute mathematical structure "just like this".[14]

If this is true, then the naturalness solution is in fact solution (C2′), where 'phenomenal properties' are replaced by 'natural properties'. The reasons for rejecting this proposal in favor of (C2‴) apply mutatis mutandis to natural properties. As a result, the naturalness solution might not get around the introduction of the technical terms in (D2′). It might simply amount to (D2′) formulated with (higher-order) natural properties. An important change in cases where inflationary conceptions of properties are involved, but otherwise not substantially different.

8. Conclusion

We have considered how the Newman problem applies to consciousness science, and shown that it threatens to undermine structural research and structural theories that target conscious experience.

The problem resides in the particular understanding of structural claims that is presumed when discussing phenomenal spaces, quality spaces, qualia spaces, experience spaces and the like. If unresolved, research that subsumes this understanding is inherently limited and prone to errors. As far as theoretical work is concerned, use of such spaces simply doesn't make sense with the usual subsumption of structural claims.

However, when one adopts are more careful definition of structural claims, the Newman problem ceases to apply. The upshot of our discussion, framed in terms of phenomenal properties for simplicity, is that if structural

[14] There is also a worry of circularity here, if in order to be able say that the world has some mathematical structure, one needs to be able to say that natural kinds have such structure.

claims like "the structure of conscious experience is E" are taken to be true if and only if the following two conditions hold, the Newman problem ceases to apply.

(C1) The elements of the domain C of are phenomenal properties of conscious experiences,

(C2) The relation R as specified by E exists, and there is a corresponding higher-order phenomenal R-property,

Here, a phenomenal property p is an R-property iff any variation that preserves the relation R preserves the phenomenal property p.

What distinguished our proposal from previous approaches is only the inclusion of R-properties in (C2). This suffices to resolve the Newman problem and the negative consequences that otherwise apply. While abstract at first, this proposal is straightforwardly applied to existing cases, and in fact builds on previous definitions of quality spaces, as explained in (Kleiner & Ludwig, 2024).

For readers with a broader background in philosophy, we have presented our proposal in general, consciousness-independent terms in Section 6 and Appendix A. Whether or not this proposal is helpful in the general discussion of the Newman problem, and whether it can be applied to domains other than consciousness, is an open question.

Acknowledgments. I would like to thank Tim Ludwig and Andrew Lee for discussions on the topic and feedback on earlier versions of the manuscript.

References

Albantakis, L., Barbosa, L., Findlay, G., Grasso, M., Haun, A. M., Marshall, W., ... others (2023). Integrated information theory 4.0: formulating the properties of phenomenal existence in physical terms. *PLoS Computational Biology*, *19*(10), e1011465.

Barrett, T. W., Manchak, J., & Weatherall, J. O. (2023). On automorphism criteria for comparing amounts of mathematical structure. *Synthese*, *201*(6), 191.

Brüntrup, G. (2011). Panpsychism and structural realism. In M. Blamauer (Ed.), *The mental as fundamental: New perspectives on panpsychism* (pp. 15–35). Ontos Verlag.

Chalmers, D. J. (2022, 07). Carnap's Second Aufbau and David Lewis's Aufbau. In *Perspectives on the Philosophy of David K. Lewis.* Oxford University Press.

Chalmers, D. J. (2023). *Phenomenal structuralism.* Talk presented at the Structuralism in Consciousness Studies Workshop at the Charité Berlin.

Chalmers, D. J., & McQueen, K. J. (2022). Consciousness and the collapse of the wave function. In S. Gao (Ed.), *Consciousness and quantum mechanics.* Oxford University Press.

Clark, A. (1993). *Sensory qualities.* Clarendon Library of Logic and Philosophy.

Fink, S. B., Kob, L., & Lyre, H. (2021). A structural constraint on neural correlates of consciousness. *Philosophy and the Mind Sciences*, *2*.

Frigg, R., & Votsis, I. (2011). Everything you always wanted to know about structural realism but were afraid to ask. *European journal for philosophy of science*, *1*, 227–276.

Kanai, R., & Fujisawa, I. (2023). Towards a universal theory of consciousness. *PsyArXiv*.

Kawakita, G., Zeleznikow-Johnston, A., Takeda, K., Tsuchiya, N., & Oizumi, M. (2023). Is my red your red?: Unsupervised alignment of qualia structures via optimal transport. *PsyArXiv*.

Ketland, J. (2004). Empirical adequacy and ramsification. *British Journal for the Philosophy of Science*, *55*(2).

Kleiner, J. (2020). Brain states matter. A reply to the unfolding argument. *Consciousness and Cognition*, *85*, 102981.

Kleiner, J. (2024). Towards a structural turn in consciousness science. *Consciousness and Cognition*, *119*, 103653.

Kleiner, J., & Hoel, E. (2021). Falsification and consciousness. *Neuroscience of Consciousness*, *2021*(1), niab001.

Kleiner, J., & Ludwig, T. (2024). What is a mathematical structure of conscious experience? *Synthese*, *203*.

Kleiner, J., & Tull, S. (2021). The mathematical structure of Integrated Information Theory. *Frontiers in Applied Mathematics and Statistics*, *6*, 74.

Lee, A. Y. (2021). Modeling mental qualities. *Philosophical Review*, *130*(2), 263–298.

Lewis, D. (1970). How to define theoretical terms. *The Journal of Philosophy*, *67*(13), 427–446.

Lewis, D. (1983). New work for a theory of universals. *Australasian journal of philosophy*, *61*(4), 343–377.

Lewis, D. K. (1986). *On the plurality of worlds.* Wiley-Blackwell.

Lyre, H. (2022). Neurophenomenal Structuralism. A philosophical agenda for a structuralist neuroscience of consciousness. *Neuroscience of Consciousness*, *2022*(1), niac012.

Mason, J. W. (2021). Model unity and the unity of consciousness: Developments in expected float entropy minimisation. *Entropy*, *23*(11), 1444.

Melia, J., & Saatsi, J. (2006). Ramseyfication and theoretical content. *The British Journal for the Philosophy of Science*.

Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review*.

Newman, M. H. (1928). Mr. Russell's "Causal Theory of Perception". *Mind*, *37*(146), 137–148.

Nida-Rümelin, M. (2018). The experience property frame work: a misleading paradigm. *Synthese*, *195*, 3361–3387.

Oizumi, M., Albantakis, L., & Tononi, G. (2014). From the phenomenology to the mechanisms of consciousness: Integrated Information Theory 3.0. *PLoS Computational Biology*, *10*(5), e1003588.

Psillos, S. (2005). *Scientific realism: How science tracks truth.* Routledge.

Rosenthal, D. (2010). How to think about mental qualities. *Philosophical Issues*, *20*, 368–393.

Rosenthal, D. (2015). Quality spaces and sensory modalities. In P. Coates & S. Coleman (Eds.), *Phenomenal qualities: sense, perception, and consciousness* (pp. 33–65). Oxford University Press Oxford, UK.

Russell, B. (2014). *The autobiography of Bertrand Russell.* Routledge.

Seager, W. (2006). The 'intrinsic nature' argument for panpsychism. *Journal of Consciousness Studies*, *13*(10-11), 129–145.

Stanley, R. P. (1999). Qualia space. *Journal of Consciousness Studies*, *6*(1), 49–60.

Tallon-Baudry, C. (2022). The topological space of subjective experience. *Trends in Cognitive Sciences*.

Tononi, G. (2015). Integrated Information Theory. *Scholarpedia*, *10*(1), 4164.

Tsuchiya, N., Phillips, S., & Saigo, H. (2022). Enriched category as a model of qualia structure based on similarity judgements. *Consciousness and Cognition*, *101*, 103319.

Wittgenstein, L. (1953). *Philosophical investigations.* Macmillan Publishing Co., New York. (translated by G. E. M. Anscombe)

Zaidi, Q., Victor, J., McDermott, J., Geffen, M., Bensmaia, S., & Cleland, T. A. (2013). Perceptual spaces: mathematical structures to neural mechanisms. *Journal of Neuroscience*, *33*(45), 17597–17602.

Appendix A. Full Definitions

Here, we provide formal details of our proposal of the general Newman problem as presented in Section 6. We denote a mathematical structure by $\mathbb{S}$. It is a tuple

$$\mathbb{S} = \big((\mathcal{A}_i)_{i\in I}, (S_j)_{j\in J}\big)$$

of domains $\mathcal{A}_i$ and functions or relations S_j. We denote the class of nomologically possible worlds by W and the properties of a world $w \in W$ by $\mathcal{A}(w)$. For expository reasons, we define the class of all properties of all worlds by

$$\mathcal{A} = \bigcup_{w\in W} A(w)\,. \tag{A.1}$$

A *variation* of a world w changes w into another world w'. Because worlds have structure, there may be various different ways to go from w to w'.[15] Therefore, in addition to specifying w and w', a variation is a partial mapping

$$v : A(w) \to A(w')\,.$$

This mapping describes how properties of the world w are replaced or reshuffled by the variation. A mapping which is not surjective, meaning that it does not map to all properties in $A(w')$, makes room for appearance of new properties of w'. A mapping which is partial, meaning that it does not specify a target for

[15] To illustrate this point, consider the following example, provided in (Kleiner & Ludwig, 2024). Let v and v' be mappings that map the numbers 1, 2, and 3 to the numbers 2, 4, and 6. The mapping v is the multiplication of every number by 2, meaning that we have $v(1) = 2$, $v(2) = 4$, $v(3) = 6$. The mapping v', on the other hand, is defined by $v(1) = 6$, $v(2) = 2$, $v(3) = 4$. If we only cared about the *sets* of elements that these mappings connect, the mappings would be equivalent: there is no difference between the set $\{2, 4, 6\}$, which is the image of v, and $\{6, 2, 4\}$, which is the image of v'. If, however, we care about the *structure* of the elements of the sets–in this case, the *ordering* of numbers–, then there is a difference. While $2 \le 4 \le 6$, it is not the case that $6 \le 2 \le 4$. Because we care about the order of the elements, we need to say which element goes where.

every property in $A(w)$, makes room for properties to disappear.

Higher-order properties are properties that are instantiated relative to other properties. If a property a requires other properties for its instantiation, we will say that the aspect *a is instantiated relative to* properties $b_1, ..., b_m$, or simply that *a is relative to* $b_1, ..., b_m$. Higher-order properties are the building blocks for our proposal to define structual claims like (2.1).

Def 3. A variation $v : A(w) \to A(w')$ **does not preserve a property** $a \in A(w)$ *relative to* $b_1, ..., b_m \in A(w)$ if and only if a is instantiated relative to $b_1, ..., b_m$ in $A(w)$, but a is not instantiated relative to $v(b_1), ..., v(b_m)$ in $A(w')$.[16]

In the case where $a \in A(w)$ is not a higher-order property, this definition reduces to the simple condition that $a \in A(w)$ but $a \notin A(w')$. The negation of the definition is also as intuitively expected: the property is present both in the source and in the target.[17]

For applications it is important to understand that this definition can fail to apply in two ways. First, it can fail because there is no a in $A(w')$ which is instantiated relative to $v(b_1), ..., v(b_m)$. This, in turn, can be the case either because there is no a in $A(w')$ at all, or because there is an a in $A(w')$ but it is instantiated relative to other aspects. Second, it can fail because one or more of the $v(b_1), ..., v(b_m)$ do not exist. The second case is possible because v is a *partial* mapping, which means aspects can disappear.

We use the term *relata* to designate those elements of a domain that are related by a structure. In the case where S is a relation R on a domain $\mathcal{A}$ and has arity m, these are the elements of the m-tuples $(b_1, ..., b_m) \in R$. In the case where S is a function $f : \mathcal{A}_1 \times ... \times \mathcal{A}_{m-1} \to \mathcal{A}_m$, the relata are the elements of the m-tuples $(b_1, ..., b_{m-1}, b_m)$ where $b_m = f(b_1, ..., b_{m-1})$, and where the other b_i range over their whole domains. For notational simplicity, we write $b_1, ..., b_m$ instead of $(b_1, ..., b_m)$ when designating relata in what follows.

Def 4. A variation $v : A(w) \to A(w')$ **preserves a structure** S *with respect to relata* $b_1, ..., b_m \in A(w)$ if and only if we have

(P1) $R\big(b_1, ..., b_m\big) = R\big(v(b_1), ..., v(b_m)\big)$ if S is a relation R, or[18]

(P2) $v\big(f(b_1, ..., b_{m-1})\big) = f\big(v(b_1), ..., v(b_{m-1})\big)$ if S is a function f.

As in the previous case, the negation of this definition is exactly what is intuitively expected: a variation does not preserve the structure if and only if the structure is satisfied before the variation, but not satisfied after the variation.[19]

For applications it is again important to see that the definition can fail for two reasons. First, it could be the case that one or more of the $v(b_i)$ do not exist in $A(e')$, if the corresponding aspect disappears. Second, the identities may fail to hold.

Def 5. A property $a \in \mathcal{A}$ **is a** S**-property** if and only if the following condition holds: A variation *does not preserve* S with respect to relata $b_1, ..., b_m$ *if and only if* the variation *does not preserve* a relative to $b_1, ..., b_m$.

This condition needs to hold true for all variations and all relata. This means that it needs to hold true for all variations of all worlds w in the class W that instantiate relata of the

[16] In (Kleiner & Ludwig, 2024), we use the term 'changes' rather than 'does not preserve'. In hindsight, we think it is easier to speak of preservation too in this case.

[17] Because the definiendum already includes the first part of the condition, the negation is as follows: A variation $v : A(w) \to A(w')$ *preserves a property* $a \in A(w)$ *relative to* $b_1, ..., b_m \in A(w)$ if and only if a is instantiated relative to $b_1, ..., b_m$ in $A(w)$ and a is also instantiated relative to $v(b_1), ..., v(b_m)$ in $A(w')$.

[18] For notational simplicity, we write $R\big(b_1, ..., b_m\big) = R\big(v(b_1), ..., v(b_m)\big)$ instead of $R\big(b_1, ..., b_m\big) \Leftrightarrow R\big(v(b_1), ..., v(b_m)\big)$.

[19] *A variation* $v : A(w) \to A(w')$ *does not preserve a structure* S *with respect to relata* $b_1, ..., b_m \in A(w)$ if and only if we have $R\big(b_1, ..., b_m\big) \neq R\big(v(b_1), ..., v(b_m)\big)$ if S is a relation R, or $v\big(f(b_1, ..., b_{m-1}\big) \neq f\big(v(b_1), ..., v(b_{m-1})\big)$ if S is a function f.
This negation agrees with the intuition because the definiendum already states part of the condition that follows, namely that $b_1, ..., b_m$ are relata of the structure S in $A(w)$, which implies that $(b_1, ..., b_m) \in R$ if S is a relation and that $f(b_1, ..., b_{m-1})$ exists in $A(w)$ if S is a function, meaning that the structure is satisfied before the variation.

structure S. Definitions 3 to 5 allow us to define Conditions (D1′) and (D2′) in more detail.

Def 6. A mathematical structure $\mathbb{S}$ is a mathematical structure of the world if and only if the following two conditions hold:

(D1′) The domains $\mathcal{A}_i$ of $\mathbb{S}$ are subsets of $\mathcal{A}$.
(D2′) For every S_j, there is a S_j-aspect in $\mathcal{A}$.

Here, $\mathcal{A}$ denotes the set of all properties of the worlds in W as defined in (A.1).

Appendix B. Objections

In this appendix, we provide the Lemmas that underlie the explanations in Section 7.

Lemma 1. If a property a is a S-property, every world that instantiates relata of S needs to instantiate a relative to these relata.

Proof. Let a be a S-property and w be any world that instantiates relata $b_1, ..., b_m$ of S. Definition 5 holds true for all variations of all worlds that instantiate relata of the structure S. Because w instantiates relata, Definition 5 applies to any variation that maps from w to any other world. Let v be any such variation. This variation either preserves S with respect to relata $b_1, ..., b_m$, or it does not preserve S with respect to relata $b_1, ..., b_m$.

Because a is a S-property, if v preserves S with respect to relata $b_1, ..., b_m$, then it preserves a relative to $b_1, ..., b_m$. But according to Definition 3, this can only be true if $a \in \mathcal{A}(w)$ relative to $b_1, ..., b_m$ (cf. Footnote 17 for details). If, on the other hand, v does not preserve S with respect to relata $b_1, ..., b_m$, then it does not preserve a relative to $b_1, ..., b_m$. But according to Definition 3, this too can only be true if $a \in \mathcal{A}(w)$ relative to $b_1, ..., b_m$. Thus both cases imply $a \in \mathcal{A}(w)$ relative to $b_1, ..., b_m$. Thus the result follows. □

The condition that corresponds to the automorphism group of a structure being trivial in the full formal setting of our definition introduced in Appendix A is that no variation of a structure, other than the identity, preserves this structure. For this case, we have the following lemma.

Let S be a structure over a domain $\mathcal{A}_0$, and assume that S contains at least two sets of relata that are properties of least two worlds.

Lemma 2. If no variation preserves S with respect to any of its relata, no S-property exists.

Proof. Let w_1 and w_2 be worlds that instantiate the relata of S. Lemma 1 implies that if there is a S-property a, both of these worlds need to instantiate a relative to the relata that they instantiate. Consider now a variation from w_1 to w_2 which maps the relata instantiated in w_1 to the relata instantiated in w_2. According to Definition 3, this variation preserves a relative to the relata instantiated in w_1 (cf. Footnote 17). Thus there is a variation that preserves a relative to said relata. If a is a S-property, Definition 5 furthermore implies that the variation preserves S with respect to those relata. This contradicts the antecedent of the claim in the Lemma. Hence no S-property can exist. □