




# What is a mathematical structure of conscious experience?

Johannes Kleiner<sup>1,2,3,4</sup>  · Tim Ludwig<sup>5</sup>

Received: 20 February 2023 / Accepted: 18 January 2024  
© The Author(s) 2024

## Abstract

Several promising approaches have been developed to represent conscious experience in terms of mathematical spaces and structures. What is missing, however, is an explicit definition of what a ‘mathematical structure of conscious experience’ is. Here, we propose such a definition. This definition provides a link between the abstract formal entities of mathematics and the concreta of conscious experience; it complements recent approaches that study quality spaces, qualia spaces, or phenomenal spaces; and it provides a general method to identify and investigate structures of conscious experience. We hope that ultimately this work provides a basis for developing a common formal language to study consciousness.

**Keywords** Quality spaces · Qualia spaces · Phenomenal spaces · Perceptual spaces · Q-spaces · Structuralism

Attempts to represent conscious experiences mathematically go back at least to 1860 (Fechner, 1860), and a large number of approaches have been developed since. They span psychophysics, philosophy, phenomenology, neuroscience, theories of consciousness, and mathematical consciousness science (Clark, 1993, 2000; Coninx, 2022; Fortier-Davy & Millièrre, 2020; Gert, 2017; Grindrod, 2018; Haun & Tononi, 2019; Hoffman & Prakash, 2014; Kleiner, 2020; Klineciewicz, 2011; Kostic, 2012; Kuehni & Schwarz, 2008; Lee, 2021, 2022; Mason, 2013; Oizumi et al., 2014; Prent-

---

✉ Johannes Kleiner  
johannes.kleiner@lmu.de

- <sup>1</sup> Munich Center for Mathematical Philosophy, Ludwig-Maximilians-Universität München, Geschwister-Scholl-Platz 1, 80539 Munich, Germany
- <sup>2</sup> Graduate School of Systemic Neurosciences, Ludwig-Maximilians-Universität München, Großhaderner Str. 2, 82152 Planegg-Martinsried, Germany
- <sup>3</sup> Institute for Psychology, University of Bamberg, Markusplatz 3, 96047 Bamberg, Germany
- <sup>4</sup> Association for Mathematical Consciousness Science, Geschwister-Scholl-Platz 1, 80539 Munich, Germany
- <sup>5</sup> Institute for Theoretical Physics, Utrecht University, Princetonplein 5, 3584 CC Utrecht, The Netherlands

ner, 2019; Renero, 2014; Resende, 2022; Rosenthal, 2010, 2015, 2016; Stanley, 1999; Tallon-Baudry, 2022; Tsuchiya & Saigo, 2021; Tsuchiya et al., 2022; Yoshimi, 2007; Young et al., 2014; Zaidi et al., 2013) and are known under various different names, including quality spaces (Clark, 1993), qualia spaces (Stanley, 1999), experience spaces (Kleiner & Hoel, 2021; Kleiner & Tull, 2021), Q-spaces (Chalmers & McQueen, *in press*), Q-structure (Lyre, 2022),  $\Phi$ -structures (Tononi, 2015), perceptual spaces (Zaidi et al., 2013), phenomenal spaces (Fink et al., 2021), spaces of subjective experience (Tallon-Baudry, 2022), and spaces of states of conscious experiences (Kleiner, 2020). The mathematical structures and spaces introduced by these approaches have enabled significant advancements in their respective fields. Nevertheless, this research remains largely fragmented. The various approaches employ different formalizations and different mathematical structures, and they presume a different, and sometimes partial, understanding of the concept of a mathematical structure or space when applied to conscious experience. What is missing, from our perspective, is a definition of the term ‘mathematical structure of conscious experience’ that clarifies how this term can and should be used.

In this article, we propose a definition of mathematical structures of conscious experience. Our main desideratum is that for a mathematical structure to be *of* conscious experience, there must be something *in* conscious experience that corresponds to that structure: a specific structural aspect of conscious experience.

Our key idea is to use variations to identify and investigate these structural aspects of conscious experience. That is because variations can serve as a binding link between conscious experiences and mathematical structures: on the one hand, variations relate to conscious experiences, because variations change aspects of conscious experiences (like qualia, qualities, or phenomenal properties); on the other hand, variations relate to mathematical structures, because they may or may not preserve them.

In defining a mathematical structure of conscious experience, our proposal does not answer the question of what this mathematical structure actually is, or which type it has. Instead, our proposal identifies the analysandum for future work on spaces and structures of conscious experience, based on which phenomenal spaces, quality spaces, qualia spaces,  $\Phi$ -structures, as well as several other related concepts, can be constructed and investigated.

This paper is structured as follows. In Sect. 1, we discuss how recent approaches relate mathematical structures to conscious experience and identify three key issues in these approaches. In Sect. 2, we present our proposal together with the necessary background information. In Sects. 3, and 4, we consider two important examples: relative similarity and topological spaces. In Sect. 5, we show how our proposal resolves the three problems identified in Sect. 1. Finally, our conclusion follows in Sect. 6.

## 1 The status quo

So where do things stand? Most of the early work that has attributed mathematical structure to conscious experience was grounded in intuition. Whether or not a specific mathematical structure is a mathematical structure of conscious experience—a structure which “pertains to”, or “belongs to” consciousness, that is—was not assessed

systematically; instead, it was assessed based on an intuitive insight of appropriateness. More recent approaches have realized the need for a more systematic method, for example Gert (2017); Lee (2021, 2023); Prentner (2019); Resende (2022); Rosenthal (2015, 2016). In this section, we analyze what we take to be the condition that underlies these approaches: a condition that justifies prescribing a mathematical structure to conscious experience. As we will see, this condition is quite natural. But, as we will demonstrate, it cannot be understood as a sufficient condition.

In a nutshell, a mathematical structure consists of two building blocks; for a detailed introduction, see Sect. 2.2. The first building block consists of one or more sets called the *domains* of the structure. The second building block are *relations or functions* which are defined on the domains. For reasons explained below, we will denote them as *structures* in the narrow sense of the term. A metric space, for example, is a mathematical structure that is defined on two domains: a set of points and the real numbers. Furthermore, it comprises a function—the so-called metric function—which maps two points to a real number. A topological space, to give another example, is a mathematical structure that is defined on a single domain: a set of points. Furthermore, it comprises a collection of unary relations, which are subsets of the domain.<sup>1</sup>

Usually, a mathematical structure also comes with *axioms*. The axioms establish conditions that the functions or relations have to satisfy. In the case of a metric structure, the axioms require the metric function to satisfy three conditions, called positive definiteness, symmetry, and triangle inequality. In the case of a topological structure, the axioms ensure that the collection includes the empty set and the whole domain, that it is closed under finite intersections, and that it is closed under arbitrary unions.

When put in these terms, recent proposals that go beyond intuitive assessments, make use, either directly or indirectly, of the following condition to justify that a specific mathematical structure is a mathematical structure of consciousness. Here, we use the term *aspect* as a placeholder for qualia, qualities, (instantiated) phenomenal properties, or similar concepts.<sup>2</sup>

**(MDC) A mathematical structure is a mathematical structure of conscious experience** if and only if the following two conditions are satisfied:

- (D1) The domains of the structure are sets whose elements correspond to aspects of conscious experiences.
- (D2) The axioms of the structure are satisfied.

In the case of the metric structure introduced in Clark (1993), for example, (D1) is satisfied because the set of points corresponds to qualities of conscious experience. The real numbers might have a phenomenal interpretation as describing degrees of similarity, as for example in Lee (2021). Condition (D2) requires positive definiteness, symmetry, and the triangle inequality to hold. This includes, for example, the condition

<sup>1</sup> A unary relation on a domain, in the mathematical sense, is a subset of the domain; see Sect. 4.

<sup>2</sup> We use the term ‘aspect’ as a placeholder for these terms because the above condition is not unanimously framed in either of these terms, and because our proposal in Sect. 2 is applicable with respect to any of these choices. In short, our goal is not to pick any one of these concepts but to offer a definition that works with respect to all of these concepts. Which concept is best suited for a particular task or domain is a philosophical question that can be answered independently of our proposal.

that “points should have distance zero just in case the qualities represented by those points are phenomenally identical” (Lee, 2021, p. 14). In the case of the topological structure introduced in Stanley (1999), to give another example, (D1) is satisfied because the domain of the structure refers to qualia. Condition (D2) would require, then, that the chosen collection of subsets satisfies the axioms of a topological space.

Prima facie, (MDC) could be taken to define what a mathematical structure of conscious experience is. However, if understood as sufficient condition, the following three problems arise.

### Problem 1: Incompatible structures

A first reason why (MDC) cannot be a sufficient condition to assess whether a mathematical structure is a mathematical structure of consciousness is that it allows for incompatible structures.

Consider, as an example, the case of topology. A basic question in topology is whether a target domain is discrete or not. A target domain is discrete if and only if its topology contains all subsets of the domain (Joshi, 1983). Otherwise, the target domain is not discrete. These two cases are exclusive, meaning that discrete and non-discrete topological structures are incompatible.

According to (MDC), conscious experience has a discrete structure. That is because any set whatsoever can be equipped with the discrete topology. Therefore, picking a set  $X$  of aspects (qualia, qualities, phenomenal properties, etc.) and choosing its discrete topology provides a mathematical structure that satisfies both conditions (D1) and (D2). But, according to (MDC), consciousness also has a non-discrete structure. That is because any set can also be equipped with a non-discrete topology. We can, for example, take an arbitrary decomposition of the set  $X$  into two subsets  $A$  and  $A^\perp$ , where  $A^\perp$  is the complement of  $A$ , and consider the topology  $\{\emptyset, A, A^\perp, X\}$ . This choice satisfies all axioms of a topology, and therefore satisfies (D2). Furthermore, it is built on the same set  $X$  as the discrete topology above, which implies that it also satisfies (D1). Therefore, the discrete and the non-discrete topological structures are both structures of conscious experience, according to (MDC).

This example shows that, if understood as a sufficient condition, (MDC) implies that two incompatible structures are both structures of conscious experience, and that they do so with respect to the exact same domain of aspects. The condition fails to determine which of the two incompatible structures is the right one.

### Problem 2: Arbitrary re-definitions

A second reason why (MDC) cannot be a sufficient condition is that it allows for arbitrary re-definitions: if one structure is given that satisfies (MDC), then any arbitrary definition of a new structure in terms of the given structure also satisfies (MDC), so long as the domains of the structure remain unchanged. If the former pertains to consciousness, so does the latter.

A simple example of this is given by rescaling a metric function. Let us suppose that  $(M, d)$  is a metric structure which pertains to consciousness according to (MDC),

where  $M$  is a set of aspects and  $d$  is the metric function, which provides a real number  $d(a, b)$  for every two aspects  $a$  and  $b$ . Since  $(M, d)$  satisfies (MDC), so does every structure  $(M, C \cdot d)$ , where  $C \cdot d$  is the multiplication of the function  $d$  by a positive real number  $C$ . Here, the number  $C$  can be chosen arbitrarily. Therefore, if one metric structure pertains to consciousness according (MDC), so does an uncountably infinite number of metric structures.

What is more, when re-defining structures, one is free to change the axioms as one pleases. For example, we could pick any function  $f$  that maps  $M$  to the positive real numbers and define a new distance function by  $(f(a) + f(b)) \cdot d(a, b)$ . This might not be a metric structure anymore, because the triangle inequality axiom might not hold. But it still satisfies positive definiteness and symmetry, and therefore satisfies (MDC), with a new set of axioms. One could even break asymmetry to get a distance function like the one applied by IIT (Kleiner & Tull, 2021). More severe cases appear with more complicated structures.

This is a problem, not only because of the unlimited number of structures that appear, but also because there is an arbitrariness in the definition of a new structure, specifically concerning the axioms. It seems strange that the axioms can be redefined at will, so as to always satisfy Condition (D2). Something is missing that restricts this arbitrariness in (MDC).

### Problem 3: Indifference to consciousness

The third reason, which speaks against the sufficiency of (MDC), is that the proposed condition seems somewhat indifferent to details of conscious experience.

To illustrate this indifference, let us consider again the discrete and non-discrete topological structures from above. As we have shown, these structures pertain to conscious experience according to (MDC). Yet, nothing more than a few lines needed to be said to establish this fact. In particular, we did not need to use any noteworthy input related to consciousness other than picking some set of aspects; and it didn't matter which aspects we picked.

It is a red flag if so short an analysis, which does not depend on consciousness in a meaningful way, establishes facts about the mathematical structure of conscious experience. The example exposes an indifference of (MDC) to details of conscious experience: the definition only relates to the different aspects, but not to the sort of mathematical object that connects these different aspects. Speaking somewhat vaguely, (MDC) does not refer to the "way" in which the different aspects of consciousness are related. This is why, in the case of topology, it allows one to draw conclusion without any noteworthy input from actual experience. This constitutes another reason that condition (MDC) is missing some important component, if used as sufficient condition.

### Cause of these problems

These three problems arise because (MDC) is not only a necessary, but also a sufficient condition: it contains an 'if' condition in addition to the 'only if' condition. In the first

problem, we showed that two incompatible mathematical structures—a discrete and a non-discrete topology—each satisfy (D1) and (D2). Because (MDC) is a sufficient condition, it follows that both structures are structures of conscious experience, according to (MDC). In light of the incompatibility of discrete and non-discrete topologies, this constitutes an issue of the definition. In the second problem, we showed that for any given structure or space that satisfies (D1) and (D2), any arbitrary redefinition yields a structure or space which also satisfies (D1) and (D2), for a suitably adapted set of axioms. Because (MDC) is a sufficient condition, this implies that the arbitrarily redefined structure is also a mathematical structure of conscious experience, which for reasons explained above, constitutes an issue as well. The third problem, finally, built on the first one and makes use of the sufficient condition in exactly the same way. Because there is no condition in (MDC) that relates to structure in the narrow sense of the term—no condition that relates to relations or functions, that is—, and because of the sufficient condition in (MDC), structures of conscious experience can be established without reference to structure in the narrow sense of the term.

## The way forward

To resolve the three problems, our task is to propose a definition for a mathematical structure of conscious experience that makes sense as a necessary and sufficient condition. This will be the content of Sect. 2.

Two desiderata guide our search. First, as is the case with (MDC), an improved definition should be *about* conscious experience in the sense that it targets qualities, qualia, instantiated phenomenal properties, or similar aspects of conscious experience, as in (D1) above. Second, there should be something *in* conscious experience—a quality, or quale or phenomenal property—that relates to structure in the narrow sense of the term. This “something” should make sure that the definition is not indifferent to conscious experience in the sense of Problem 3, and that the definition refers to functions or relations in a meaningful way, so as to stop arbitrary re-definitions (Problem 2). The proposal which we present in the next section is the result of our search.

Despite the above-mentioned problems, we think that (MDC) is an important condition. It might not be suitable as a sufficient condition, but it is valuable as a necessary condition. If one understands mathematics pragmatically as constituting a *language*—a body of symbols and terms with rules that connect these—, then mathematics can be used to describe phenomena, much like the English language can. Looking back at Condition (MDC) after our analysis, and presuming this pragmatic conception of mathematics, we think that (MDC) is best understood as an expression of what it takes for a mathematical structure to *describe* conscious experience. That is, (MDC) might be a valuable descriptive tool that utilizes mathematical structure to represent information on how aspects are related to each other (as explicated by (D1) and (D2)).

Because of this, we will refer to a mathematical structure that satisfies (MDC) as a mathematical structure that ‘describes conscious experience’ in what follows. The new condition which we develop below contains (MDC) as necessary part; this is aligned with the intuition that any mathematical structure of conscious experience also describes conscious experience.

## 2 Mathematical structures of conscious experience

In this section, we provide a definition of what mathematical structures of conscious experience are. Based on this definition, phenomenal spaces, quality spaces, qualia spaces, and related structures can be constructed and investigated. The definition embodies a way to think and work with mathematical structures when applied to conscious experience.

Our key desideratum in improving (MDC), explained above, is that for a mathematical structure to be a mathematical structure of conscious experience, rather than just a descriptive tool for conscious experience, there must be a structural aspect in conscious experience that relates to that structure. A major goal of this section is to explain this in detail. Denoting a mathematical structure by  $S$ , we call this structural aspect an  $S$ -aspect.

To make sense of what an  $S$ -aspect is, we need to understand how aspects (like qualia, qualities, or phenomenal properties) relate to mathematical structures. While aspects may have an arity, meaning they may be instantiated relative to other aspects, they are not experienced as having a mathematical structure per se (unless, of course, they are aspects of experiences of mathematical structures themselves, such as of geometric shapes). Therefore, relating aspects to mathematical structures requires a tool that applies to both: concrete aspects of conscious experience and abstract formal entities. Variations provide such a tool.

In general, a variation is a change of something into something else; in our case, it is a change of one experience into another experience. Such variations may be induced by external stimuli or interventions, occur naturally, or be subjected to imagination ('imaginary variations' (Husserl, 1936/1970)). Variations are directly related to aspects of conscious experiences because a variation can change an aspect. This is the case iff an aspect is part of the experience before the variation but isn't part of the experience after the variation. And variations are also intimately related to mathematical structures, because they may or may not preserve them, as explained in detail below. An  $S$ -aspect, then, is an aspect that is changed by a variation if and only if the variation does not preserve the structure  $S$ . To explain this in detail is the purpose of the remainder of this section.

### 2.1 Terminology and notation

Here, we introduce the key terms we use to define mathematical structures of conscious experience. These terms are *conscious experiences*, *aspects* of conscious experiences, and *variations* of conscious experiences. The introduction proceeds axiomatically, so that our construction does not rely on a specific choice of these concepts. Rather, any choice of these concepts that is compatible with the requirements below can be the basis of an application of our definition.

Our construction is based on a set  $E$  of *conscious experiences* of an experiencing subject. We denote individual conscious experiences in that set by symbols like  $e$  and  $e'$ ; formally  $e, e' \in E$ . From a theoretical or philosophical perspective, one may think of the set  $E$  as comprising all conscious experiences which one experiencing

subject can have, i.e. all nomologically possible experiences of that subject. From an experimental or phenomenological perspective, one may think of this set as comprising all conscious experiences that can be induced in the lab or in introspection. Different such choices may lead to different mathematical structures being accessible.

We use the term *aspect* as a placeholder for concepts such as *qualia* (Tye, 2021), *qualities* (Clark, 1993), *mental qualities* (Rosenthal, 2010), or (instantiated) *phenomenal properties*.<sup>3</sup> For every experience  $e \in E$ , we denote the set of aspects instantiated in this experience by  $A(e)$ . The set of all aspects of the experiences in  $E$ , denoted by  $\mathcal{A}$ , is the union of all  $A(e)$ ; formally  $\mathcal{A} = \bigcup_{e \in E} A(e)$ . Individual aspects, that are members of  $\mathcal{A}$ , will be denoted by small letters such as  $a, b, c$ . When explaining examples, we will often use the abbreviation ‘ $a$  is the experience of ...’ as a shorthand for saying ‘ $a$  is a ... aspect of an experience’. For example, ‘ $a$  is the experience of red color’ means ‘ $a$  is a red color aspect of an experience’.

Some aspects may require other aspects for their instantiation. For example, it is usually the case that an experience of relative similarity is an experience of relative similarity of something, for example two color aspects relative to a third color aspect. If an aspect  $a$  requires other aspects for its instantiation, we will say that the aspect  $a$  *is instantiated relative to* aspects  $b_1, \dots, b_m$ , or simply that  $a$  *is relative to*  $b_1, \dots, b_m$ . Aspects which are instantiated relative to other aspects are the building blocks for the structure of conscious experience.

A *variation* of a conscious experience  $e$  changes  $e$  into another experience  $e'$ . Because experiences have structure, there may be various different ways to go from  $e$  to  $e'$ .<sup>4</sup> Therefore, in addition to specifying  $e$  and  $e'$ , a variation is a partial mapping

$$v : A(e) \rightarrow A(e') .$$

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

<sup>3</sup> Many other concepts work as well. For example, if one works with an atomistic conception of states of consciousness, where the total phenomenal state of a subject—what it is like to be that subject at a particular time—is built up from individual atomic states of consciousness, one can take  $e$  to denote the total phenomenal state and aspects to be the *states of consciousness* in that total state. Another example would be to take aspects to denote phenomenal distinctions as used in Integrated Information Theory (Tononi, 2015). What matters for our definition to be applicable is only that according to one’s chosen concept of conscious experience, every conscious experience exhibits a set of aspects.

<sup>4</sup> To illustrate this point, consider, for example, the following two mappings  $v$  and  $v'$  which 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  $2 \leq 4 \leq 6$ , it is not the case that  $6 \leq 2 \leq 4$ . Because we care about the order of the elements, we need to say which element goes where.

## 2.2 What is a mathematical structure?

To find a rigorous definition of the mathematical structure of conscious experience, we need to work with a rigorous definition of mathematical structure. Mathematical logic provides this definition, which we now review.

A *mathematical structure*  $\mathbb{S}$  consists of two things: domains, on the one hand, and functions or relations, on the other hand. We now introduce these concepts based on two simple examples.

The *domains* of a structure  $\mathbb{S}$  are the sets on which the structure is built. We denote them by  $\mathcal{A}_i$ , where  $i$  is some index in a parameter range  $I$ . In the case of a metric structure, for example, the domains would be  $\mathcal{A}_1 = M$  and  $\mathcal{A}_2 = \mathbb{R}$ , where  $M$  is a set of points and  $\mathbb{R}$  denotes the real numbers, understood as a set. In the case of a strict partial order, there is just one domain  $\mathcal{A}$ , which contains the elements that are to be ordered.

The second ingredient are functions and/or relations. *Functions*  $f$  map some of the domains to other domains. In the case of a metric structure, the function would be a metric function  $d : M \times M \rightarrow \mathbb{R}$ , which maps from  $\mathcal{A}_1 \times \mathcal{A}_1$  to  $\mathcal{A}_2$ . A *relation*  $R$ , in the mathematical sense, is a subset of the  $m$ -fold product  $\mathcal{A}_i \times \cdots \times \mathcal{A}_i$ . Here,  $\mathcal{A}_i$  is the domain on which the relation is defined, and  $m$  is the arity of the relation, which expresses how many relata the relation relates. The product is usually just written as  $\mathcal{A}_i^m$ . In the case of a strict partial order, the relation is binary, which means that  $R$  is a subset of  $\mathcal{A}^2$ . For binary relations, one usually uses notation like  $a < b$  instead of writing  $(a, b) \in R$ .

In almost all cases, mathematical structures also come with *axioms*, which establish conditions that the functions or relations have to satisfy. They are useful because they constrain and classify the structure at hand. For  $\mathbb{S}$  to be a metric structure, for example, the function  $d$  has to satisfy the axioms of positive definiteness, symmetry, and triangle inequality (Rudin, 1976). For  $\mathbb{S}$  to be a strict partial order, the relation  $R$  has to be irreflexive, asymmetric, and transitive (Joshi, 1989).

To have a nice and compact notation, we will use one symbol  $S_j$  to denote both functions and relations. That is because, in any concrete proposal, it is always clear whether  $S_j$  is a function or a relation.<sup>5</sup> The index  $j$  takes values in some parameter range  $J$  that specifies how many functions or relations there are. Using this notation, we can represent the definition of mathematical structure provided by mathematical logic as follows:

**A mathematical structure  $\mathbb{S}$  is a tuple**

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

of domains  $\mathcal{A}_i$  and functions or relations  $S_j$ .

<sup>5</sup> In mathematical logic, mathematical structures are denoted as triples of domains, relations, and functions. However, in our case, using just one symbol for functions and relations improves readability substantially.

For given domains  $\mathcal{A}_i$ , the mathematical structure  $\mathbb{S}$  is fully determined by the  $S_j$ . Thus, we can also refer to  $S_j$  as ‘structures’, if the domains are clear from context. For simplicity, we can drop the index  $j$  and simply write  $S$  whenever we consider just one such structure.

As a final step in this section, we introduce the *relata* of a structure  $S$ . This will be helpful to present definitions concisely below. The term *relata* designates those elements 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  $(b_1, \dots, b_m) \in R$ . In the case where  $S$  is a function  $f : \mathcal{A}_1 \times \dots \times \mathcal{A}_{m-1} \rightarrow \mathcal{A}_m$ , the *relata* are the elements of the  $m$ -tuples  $(b_1, \dots, b_{m-1}, b_m)$  where  $b_m = f(b_1, \dots, b_{m-1})$ , and where the other  $b_i$  range over their whole domains. For notational simplicity, we write  $b_1, \dots, b_m$  instead of  $(b_1, \dots, b_m)$  when designating *relata* below.

### 2.3 What is a mathematical structure of conscious experience?

Finally, to the heart of the matter! We recall that we have so far identified two desiderata for a mathematical structure  $\mathbb{S}$  to be a mathematical structure of conscious experience. First, it should be *about* conscious experiences in the sense that its domains should correspond to aspects of conscious experiences. Second, there should be aspects *in* conscious experience that relate to the structure  $\mathbb{S}$ . The following definition satisfies these two desiderata. Its explanation is the task of the remainder of this section.

**(MSC) A mathematical structure  $\mathbb{S}$  is a mathematical structure of conscious experience** if and only if the following two conditions hold:

(S1) The domains  $\mathcal{A}_i$  of  $\mathbb{S}$  are subsets of  $\mathcal{A}$ .

(S2) For every  $S_j$ , there is an  $S_j$ -aspect in  $\mathcal{A}$ .

Here,  $\mathcal{A}$  denotes the set of all aspects of the experiences in  $E$ ; formally  $\mathcal{A} = \bigcup_{e \in E} A(e)$ , the  $\mathcal{A}_i$  denote the domains of the structure  $\mathbb{S}$ , and the  $S_j$ -aspects are defined below.

Condition (S1) guarantees that the first desideratum is satisfied. Condition (S2) guarantees that the second desideratum is satisfied. Furthermore, whenever a certain *type* of structure (metric, topological, partial order, manifold, etc.) is claimed to be a structure of conscious experience, the axioms that constrain and classify that type have to hold. Therefore, any mathematical structure of conscious experience (MSC) is also a mathematical structure that describes conscious experience according to (MDC). The condition that has been applied in previous proposals remains a necessary condition in (MSC).

The remaining task of this section is to explain what an  $S_j$ -aspect is. For notational simplicity, we use the symbol  $S$  to denote  $S_j$ . As we have emphasized before, variations are key to understand the structure of conscious experience, because they link aspects and structure. Therefore, to be able to precisely define what an  $S$ -aspect is, we need to understand how variations relate to aspects, on the one hand, and structures, on the other hand. Our strategy is to first discuss how variations relate to aspects. This amounts to specifying what precisely it means for a variation to change an aspect.

Second, we focus on how variations relate to mathematical structure. This amounts to explaining what it means for a variation to preserve a structure. Finally, combining these two steps allows us to understand  $S$ -aspects and provide a useful definition.

What does it mean for a variation  $v : A(e) \rightarrow A(e')$  to change aspects? The underlying idea is simply that an aspect is present in the source of the variation,  $A(e)$ , but not present any more in the target of the variation,  $A(e')$ . We need to take into account, though, that aspects are often instantiated relative to other aspects (see Sect. 2.1). This can be done as follows.

**A variation  $v : A(e) \rightarrow A(e')$  changes an aspect  $a \in A(e)$  relative to  $b_1, \dots, b_m \in A(e)$  if and only if  $a$  is instantiated relative to  $b_1, \dots, b_m$  in  $A(e)$ , but  $a$  is not instantiated relative to  $v(b_1), \dots, v(b_m)$  in  $A(e')$ .**

In the case where  $a \in A(e)$  is not instantiated relative to other aspects, the definition indeed reduces to the simple condition that  $a \in A(e)$  but  $a \notin A(e')$ . The negation of the definition is also as intuitively expected: the aspect is present both in the source and in the target.<sup>6</sup>

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  $A(e')$  which is instantiated relative to  $v(b_1), \dots, v(b_m)$ . This, in turn, can be the case either because there is no  $a$  in  $A(e')$  at all, or because there is an  $a$  in  $A(e')$  but it is instantiated relative to other aspects. Second, it can fail because one or more of the  $v(b_1), \dots, v(b_m)$  do not exist. The second case is possible because  $v$  is a *partial* mapping, which means aspects can disappear.

What does it mean for a variation to preserve a mathematical structure? The underlying idea is that a variation preserves the structure if and only if the structure is satisfied before the variation and remains to be satisfied after the variation. By its very nature, this is a mathematical condition, namely the condition of being a homomorphism (Mileti, 2022). The definition of a homomorphism, though, always applies to all elements of a domain at once. For our case, it is best to refine this definition to a single set of relata.<sup>7</sup>

**A variation  $v : A(e) \rightarrow A(e')$  preserves a structure  $S$  with respect to relata  $b_1, \dots, b_m \in A(e)$  if and only if we have**  
 (P1)  $R(b_1, \dots, b_m) = R(v(b_1), \dots, v(b_m))$  if  $S$  is a relation  $R$ , or  
 (P2)  $v(f(b_1, \dots, b_{m-1})) = f(v(b_1), \dots, v(b_{m-1}))$  if  $S$  is a function  $f$ .

<sup>6</sup> Because the definiendum already includes the first part of the condition, the negation is as follows: A variation  $v : A(e) \rightarrow A(e')$  does not change an aspect  $a \in A(e)$  relative to  $b_1, \dots, b_m \in A(e)$  if and only if  $a$  is instantiated relative to  $b_1, \dots, b_m$  in  $A(e)$  and  $a$  is also instantiated relative to  $v(b_1), \dots, v(b_m)$  in  $A(e')$ . We felt that is the best way of writing things to optimize clarity.

<sup>7</sup> For notational simplicity, we write  $R(b_1, \dots, b_m) = R(v(b_1), \dots, v(b_m))$  instead of  $R(b_1, \dots, b_m) \Leftrightarrow R(v(b_1), \dots, v(b_m))$ .

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.<sup>8</sup>

For applications it is again important to see that the definition can fail to apply 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.

We now have the keys to understand  $S$ -aspects. The underlying idea is that an  $S$ -aspect is an aspect that, under any variation, behaves exactly as the structure  $S$  does: whenever  $S$  is preserved, the  $S$ -aspect does not change, and whenever the  $S$ -aspect changes, the structure  $S$  is not preserved. This is expressed by the following definition.

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

Here, the condition needs to hold true for all variations and all relata. This means that it needs to hold true for all variations of all experiences  $e$  in the set  $E$  that instantiate relata of the structure  $S$ .

This concludes our proposal for the definition of the mathematical structure of conscious experience. It is a structure whose domains correspond to sets of aspects, and which contains an  $S$ -aspect for every relation or function of the structure. In the next two sections, we apply this definition to two examples. On the one hand, these examples illustrate the definition. On the other hand, they provide new insights to structures that have been featured prominently in previous approaches.

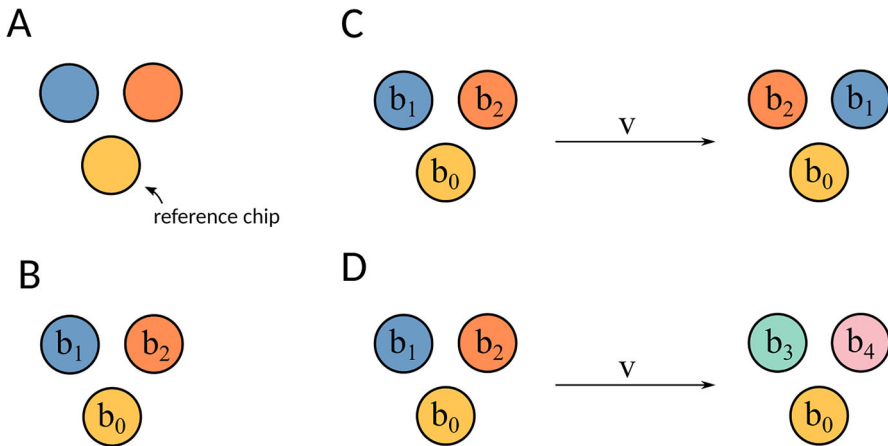
### 3 Relative similarity

Our first example concerns relative similarity, which plays an important role, for example, in the construction of quality spaces by Clark (1993, 2000).

A first step in applying our definition is to choose a set  $E$ . Here we take  $E$  to comprise experiences of three color chips, as indicated in Fig. 1A, where one of the chip (the reference) has a fixed color coating and the others vary in a range of color coatings  $\Lambda$ . A color coating is a physical stimuli.

The second step is to specify the set of aspects  $A(e)$  for every experience  $e \in E$ . Here, we take  $A(e)$  to comprise: (a) the color qualities in  $e$ , that is, the experienced colors of the individual chips; (b) positional qualities of the color experiences, that is, which chip has which color; and (c) the experience of *relative similarity*. Relative similarity is an experience of one pair of aspects to be more, less, or equally similar to

<sup>8</sup> A variation  $v : A(e) \rightarrow A(e')$  does not preserve a structure  $S$  with respect to relata  $b_1, \dots, b_m \in A(e)$  if and only if we have  $R(b_1, \dots, b_m) \neq R(v(b_1), \dots, v(b_m))$  if  $S$  is a relation  $R$ , or  $v(f(b_1, \dots, b_{m-1})) \neq f(v(b_1), \dots, v(b_{m-1}))$  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  $b_1, \dots, b_m$  are relata of the structure  $S$  in  $A(e)$ , which implies that  $(b_1, \dots, b_m) \in R$  if  $S$  is a relation and that  $f(b_1, \dots, b_{m-1})$  exists in  $A(e)$  if  $S$  is a function, meaning that the structure is satisfied before the variation.



**Fig. 1** To help explain the example of relative similarity, this figure illustrates experiences with color qualities and variations thereof. Subfigure **A** illustrates an experience of three color chips as well as the concept of *relative similarity*: many readers will experience the color of the top-left color chip to be *less similar* to the reference chip than the color of the top-right color chip. Subfigure **B** illustrates our notation for the color aspects corresponding to the color chips. Subfigures **C** and **D** illustrate variations  $v$  of experiences: a swap of two color aspects in **C**; and a replacement of two color aspects in **D**

each other than another pair of aspects; here, the two pairs have to have one aspect—the reference—in common. In Fig. 1A, for example, the color of the top left chip will, for many readers, be less similar to the reference chip than the color of the top right chip. An experience  $e$  in the set  $E$  may exhibit many other aspects as well. However,  $A(e)$  only comprises those which are relevant for the construction at hand.

To pick out relative similarity more precisely, we let  $b_0$ ,  $b_1$  and  $b_2$  denote the color aspects of the three chips in an experience  $e$ , where  $b_0$  is the color aspect of the reference; see Fig. 1B. For some experience  $e$ , it might be the case that the colors  $b_1$  and  $b_0$  are experienced as less similar to each other than the colors  $b_2$  and  $b_0$ . In this case, the experience  $e$  has a relative similarity aspect in the above sense; we denote this “less-similar” relative similarity aspect by  $a$ . So,  $a$  is an aspect of  $e$ , and it is instantiated relative to  $b_1$  and  $b_2$ . (The aspect  $a$  is also relative to  $b_0$ . But since  $b_0$  does not vary in  $E$  we can leave this implicit.)

Variations change one experience  $e$  into another experience  $e'$ . An example for a variation would be a swap of the coatings of the two non-reference chips, as in Fig. 1C. Another example for a variation would be to change the coatings of both non-reference chips to some other coating in  $\Lambda$ , as in Fig. 1D. Formally, variations are represented by mappings  $v : A(e) \rightarrow A(e')$ . In the first example, Fig. 1C, the mapping is of the form  $v(b_1) = b_2$  and  $v(b_2) = b_1$ , and  $v(c) = c$  for all other aspects  $c$ , except for the relative similarity aspect  $a$ , which is discussed in detail below. In the second example, Fig. 1D, the mapping is as in the first example but with  $v(b_1) = b_3$  and  $v(b_2) = b_4$ .

The key question of this example is: Is there a mathematical structure of conscious experience which corresponds to relative similarity? To answer this question, we propose a mathematical structure and check whether this structure satisfies (MSC).

The words “less similar than” in the description of relative similarity already indicate that some order, in the mathematical sense of the word, might be involved. For reasons that will become clear below, we propose a strict partial order as mathematical structure. Our task in the remainder of this section is to show that this proposal indeed satisfies (MSC) with respect to experienced relative similarity. A strict partial order  $(\mathcal{C}, <)$ , consists of a set  $\mathcal{C}$ , which is the domain of the structure, and a binary relation ‘ $<$ ’ on  $\mathcal{C}$ . For all  $x, y, z \in \mathcal{C}$ , this binary relation has to satisfy the following axioms:

- *Irreflexivity*, meaning that there is no  $x \in \mathcal{C}$  with  $x < x$ .
- *Asymmetry*, meaning that if  $x < y$ , then it is not the case that  $y < x$ .
- *Transitivity*, meaning that if  $x < y$  and  $y < z$ , then also  $x < z$ .

In order to turn a strict partial order into a proposal for a mathematical structure of conscious experience, we need to specify how the set  $\mathcal{C}$  and the relation  $<$  relate to aspects of conscious experience. For the set  $\mathcal{C}$  we choose the color qualities of the experiences in  $E$ , meaning that  $\mathcal{C}$  now comprises the color qualities evoked by the coatings  $\Lambda$  of the chips we consider. For example, it contains what we have labelled  $b_0, b_1, b_2, b_3$  and  $b_4$  in Fig. 1. For the relation, we define  $b_i < b_j$  if and only if  $b_i$  is experienced as less similar to  $b_0$  than  $b_j$  is to  $b_0$ . (Since relative similarity, as defined above, depends on the choice of reference  $b_0$ , it would be more precise to write  $<_{b_0}$  instead of  $<$ . However, to simplify the notation, we keep the reference implicit.)

For this proposal to make sense, we first need to check whether the axioms are satisfied. If they were not satisfied, the proposal could still be a structure of conscious experience; but it wouldn’t be a strict partial order. That’s why the axioms are not explicitly mentioned in (MSC). Irreflexivity is satisfied because no color quality is experienced as less similar to the reference than itself. Asymmetry is satisfied because if  $b_i$  is less similar to the reference than  $b_j$ , then  $b_j$  is not less similar to the reference than  $b_i$ .

The use of terms like ‘less similar to’ in natural language suggests that transitivity is also satisfied; it suggests that, if  $b_i$  is less similar to the reference than  $b_j$  and  $b_j$  is less similar to the reference than  $b_k$ , then  $b_i$  should be less similar to the reference than  $b_k$ . But it might very well be the case that natural language is not precise enough to describe its target domain. The use of natural language may be justified in simple cases, or even in a majority of cases, but whether or not transitivity holds for all  $b_i, b_j, b_k \in \mathcal{C}$  is, ultimately, an empirical question. For the purpose of this example, we’re going to assume that transitivity holds as well.

Having checked that the axioms hold—that is, that the proposal is indeed a strict partial order—we can proceed to check whether the structure is a mathematical structure of conscious experience according to (MSC). Concerning Condition (S1), there is one domain  $\mathcal{C}$  and it consists of color qualities, so this condition is satisfied. Therefore, only Condition (S2) remains to be checked.

We now show that the relative similarity aspect  $a$ , as defined above, is in fact an  $S$ -aspect, where  $S$  is the ‘ $<$ ’ relation on  $\mathcal{C}$ . That is, it is a  $<$ -aspect. To see that this is true, we have to show that a variation does not preserve  $<$  with respect to relata  $b_1$  and  $b_2$  if and only if the variation changes  $a$  relative to  $b_1$  and  $b_2$ .

Consider any variation  $v : A(e) \rightarrow A(e')$  that does not preserve  $<$  with respect to relata  $b_1, b_2 \in A(e)$ . Two aspects  $b_1$  and  $b_2$  are relata of  $<$  if either  $b_1 < b_2$  or

$b_2 < b_1$ . We focus on the first case as the other one follows from the first by renaming  $b_2$  and  $b_1$  in what follows. By definition of the  $<$  relation,  $b_1 < b_2$  means that  $b_1$  is experienced as less similar to the reference than  $b_2$ . Therefore, there is also a relative similarity aspect  $a \in A(e)$  as defined above. As explained in Sect. 2.3, there can be two ways in which the variation  $v$  might not preserve  $<$ . Either  $v(b_1)$  or  $v(b_2)$  are not defined, or, if they are defined, it is not the case that  $v(b_1) < v(b_2)$ . In the former case, there cannot be an  $a$  in  $A(e')$  relative to  $v(b_1)$  or  $v(b_2)$ , simply because the latter do not both exist. In the latter case, it follows from the definition of  $<$  that  $v(b_1)$  is not experienced as less similar to the reference than  $v(b_2)$ . So, there is no  $a \in A(e')$  relative to  $v(b_1)$  and  $v(b_2)$ . Hence, we may conclude that  $v$  changes  $a$  relative to  $b_1$  and  $b_2$ .

For the other case, let  $v : A(e) \rightarrow A(e')$  be a variation which preserves  $<$  with respect to  $b_1$  and  $b_2$ . As before, this implies that  $a$  is in  $A(e)$  relative to  $b_1$  and  $b_2$ . Because  $v$  preserves  $<$ ,  $v(b_1)$  and  $v(b_2)$  both exist and we also have  $v(b_1) < v(b_2)$ . Applying the definition of  $<$  then implies that  $a$  is also in  $A(e')$  relative to  $v(b_1)$  and  $v(b_2)$ . Hence  $v$  does not change  $a$  relative to  $b_1$  and  $b_2$ .

Because in both of these cases,  $v$  was arbitrary, it follows that  $a$  is indeed a  $<$ -aspect. Therefore, Conditions (S1) and (S2) of (MSC) are both satisfied, and the strict partial order  $(\mathcal{C}, <)$  is indeed a mathematical structure of conscious experience; it is the mathematical structure of relative similarity of color experiences with respect to  $b_0$ .

## 4 Phenomenal unity and topological structure

Our second example concerns topological structure. Interestingly, this is intimately tied to phenomenal unity, the thesis that phenomenal states of a subject at a given time are unified (Bayne & Chalmers, 2003). Phenomenal unity gives rise to a mathematical structure of conscious experience.<sup>9</sup>

Recall that we have introduced the set  $A(e)$  to denote aspects of the conscious experience  $e$ , where we have used the term ‘aspect’ as a placeholder for concepts like qualia, qualities, or (instantiated) phenomenal properties. Most examples of these concepts are “independent” from the experience in which they occur; they could be experienced together with a largely different set of aspects in a different experience. Yet, experiences seem unified; their aspects are experienced as tied together in some essential way. This raises the question of what underlies this experience of the *unity of a conscious experience*? As we will see, somewhat surprisingly, the answer is: a topological structure of conscious experience.

Much has been written about the question of phenomenal unity in the literature, for example Bayne (2012); Bayne and Chalmers (2003); Cleeremans and Frith (2003); Mason (2021); Prentner (2019); Roelofs (2016); Wiese (2018), and in order to make

<sup>9</sup> A connection between topology and phenomenal unity has already been conjectured in Prentner (2019), where an attempt was made to construct a topological space based on a binary relation that describes the “overlap” of mental objects. The construction only leads to the weaker notion of a pre-topology, but should be regarded as an important first step in this direction. For a summary of the formal construction, see Kleiner (2020, Example 3.22).

use of some of the results, we assume that the term ‘aspect’ denotes an instantiated phenomenal property or quale. The set of aspects  $A(e)$ , then, comprises the phenomenal properties or qualia which are instantiated in the experience  $e$ , also called the *phenomenal states* of the experience  $e$ .<sup>10</sup> Our question, then, is what it means that “any set of phenomenal states of a subject at a time is phenomenally unified” (Bayne & Chalmers, 2003, p. 12).

There are various answers one might give to this question. A promising answer is the so-called *subsumptive unity thesis*, developed in Bayne and Chalmers (2003):

For any set of phenomenal states of a subject at a time, the subject has a phenomenal state that subsumes each of the states in that set. (Bayne & Chalmers, 2003, p. 20)

According to this thesis, what underlies the experience of the unity of a conscious experience is that for any set  $X$  of phenomenal states in the conscious experience, there is a further phenomenal state that subsumes each of the states in  $X$ . This phenomenal state characterizes what it is like to be in all of the states of  $X$  at once (Bayne & Chalmers, 2003, p. 20).

Put in terms of aspects, the subsumptive unity thesis says that for any set  $X \subset A(e)$  of aspects of an experience, there is an additional aspect in  $A(e)$  that subsumes the aspects in  $X$ . This aspect is the experience of what it is like to experience the aspects in  $X$  as part of one experience  $e$  together; this aspect is the experience that the aspects in  $X$  are *unified*, as we will say. Let us call this aspect the *phenomenal unity aspect* of  $X$  and denote it by  $a_X$ . It is instantiated relative to the elements of  $X$ .

Phenomenal unity gives rise to a mathematical structure of conscious experience. To see how, let us use the symbol  $\mathcal{T}$  to denote a collection of subsets of  $A(e)$ , to be specified in more detail below. Every subset of  $A(e)$  is a unary relation on  $A(e)$ ,<sup>11</sup> and hence also on the set  $\mathcal{A}$  that comprises all aspects of the experiences in  $E$ . Therefore,  $(\mathcal{A}, \mathcal{T})$  is a mathematical structure; it has the domain  $\mathcal{A}$  and its structures are the unary relations in  $\mathcal{T}$ . As we show next, because of the subsumptive unity thesis, the mathematical structure  $(\mathcal{A}, \mathcal{T})$  is a mathematical structure of conscious experience according to (MSC).

Because  $\mathcal{A}$  is the set of all aspects of  $E$ , Condition (S1) of (MSC) is satisfied. Therefore, only Condition (S2) remains to be checked. This condition is satisfied because for every set  $X \in \mathcal{T}$ , the phenomenal unity aspect  $a_X$  is an  $S$ -aspect for  $S = X$ ; an  $X$ -aspect for short. To show that this is the case, we need to check that a variation does not preserve  $X$  with respect to relata  $b_1, \dots, b_m$  if and only if it changes  $a_X$  relative to  $b_1, \dots, b_m$ . Let  $v : A(e) \rightarrow A(e')$  be a variation that does not preserve  $X$  with respect to relata  $b_1, \dots, b_m$ . The relata of the subset  $X$  are the elements of that subset. Therefore, we have  $b_1, \dots, b_m \in A(e)$ , so that the subsumptive unity thesis implies that there is a phenomenal unity aspect  $a_X$  relative to the  $b_1, \dots, b_m$  in  $A(e)$ .

<sup>10</sup> A *phenomenal state* is an instantiation of a phenomenal property, or quale, by a subject at a given time. This instantiation constitutes part of the experience of the subject at the time. An experience  $e$ , in our terminology, is an experience of a subject at a given time. Hence, a phenomenal state is an instantiation of a phenomenal property, or quale, in an experience  $e$ .

<sup>11</sup> An  $m$ -ary relation on a set  $X$  is a subset  $R$  of  $X^m$ . Hence, a unary relation, where  $m = 1$ , is a subset of  $X$ .

The condition that  $v$  does not preserve  $X$  furthermore implies that either not all of the  $v(b_i)$  exist or that at least one of them is not in the set  $X$ . Therefore, there is no phenomenal unity aspect  $a_X$  relative to  $v(b_1), \dots, v(b_m)$  in  $A(e')$ . Hence, the variation  $v$  changes  $a_X$  relative to  $b_1, \dots, b_m \in X$ . Vice versa, let  $v : A(e) \rightarrow A(e')$  be a variation which preserves  $X$  with respect to  $b_1, \dots, b_m$ . This implies that  $a_X$  is instantiated relative to  $b_1, \dots, b_m$  in  $A(e)$ . The condition that  $v$  preserves  $X$  furthermore implies that  $v(b_1), \dots, v(b_m)$  exist, and that they are elements of  $X$ . Therefore,  $a_X$  is also instantiated relative to  $v(b_1), \dots, v(b_m)$  in  $A(e')$ . This shows that the variation does not change  $a_X$  relative to  $b_1, \dots, b_m$ . Thus,  $a_X$  is indeed an  $X$ -aspect. And because that is true for any  $X \in \mathcal{T}$ ,  $(\mathcal{A}, \mathcal{T})$  indeed satisfies Condition (S2) and hence (MSC).

The previous paragraph proves that, if the subsumptive unity thesis holds true for all sets  $X$  in  $\mathcal{T}$ , then  $(\mathcal{A}, \mathcal{T})$  is indeed a mathematical structure of conscious experience. As we will explain next, this structure is intimately tied to a topological structure.

A topological structure  $(M, \mathcal{T})$  consists of a set  $M$  and a collection  $\mathcal{T}$  of subsets of  $M$ . The collection has to satisfy three axioms, and there are a few different ways of formulating these axioms. Here, we choose the formulation that corresponds to what is usually called ‘closed sets’. The axioms are:

- The empty set  $\emptyset$  and the whole set  $M$  are both in  $\mathcal{T}$ .
- The intersection of any collection of sets of  $\mathcal{T}$  is also in  $\mathcal{T}$ .
- The union of any finite number of sets of  $\mathcal{T}$  is also in  $\mathcal{T}$ .

Are these axioms satisfied by the structure  $(\mathcal{A}, \mathcal{T})$  induced by phenomenal unity?

To answer this question, it is important to note that the subsumptive unity thesis does not provide a phenomenal unity aspect  $a_X$  for every subset of  $\mathcal{A}$ . It can only provide such an aspect for a set of aspects that are actually experienced together. That is, it can only provide such an aspect for a subset  $X$  of  $A(e)$ . Therefore,  $\mathcal{T}$  is not the discrete topology introduced in Sect. 1. Second, it also cannot be the case that it provides a phenomenal unity aspect for every subset of  $A(e)$ . That’s because then there would be an infinite regress: for every subset  $X$  of  $A(e)$  there would be a new aspect  $a_X$  in  $A(e)$ , giving a new subset  $X \cup \{a_X\}$  that would give a new phenomenal unity aspect  $a_{X \cup \{a_X\}}$ , and so forth. This problem is well-known in the literature (Bayne, 2005; Wiese, 2018). Rather, we take it, the quantifier ‘any set’ in the subsumptive unity thesis must be understood as ‘any set of aspects that are experienced as being unified’. While it is arguably the case that the whole set of aspects  $A(e)$  of an experience is always experienced as unified—by which we mean: the whole set of aspects is experienced—, introspection suggests that we consciously experience only a select group of aspects as unified at a time.<sup>12</sup>

So, which sets of aspects do we experience as unified? While we cannot give a general answer to this question here, there is a special case where a sufficiently detailed specification can be given: the case of regions in visual experience. Here, ‘regions’ are

<sup>12</sup> This solves the infinite regress problem because, arguably, we do not always experience the phenomenal unity aspects as unified with the sets they correspond to. So, there is not always a phenomenal unity aspect  $a_{X \cup \{a_X\}}$  for the set that consists of  $a_X$  and  $X$ .

sets of positions of the space that visually perceived objects occupy.<sup>13</sup> The positions in a region are experienced as unified. Therefore, the regions of visual experience are members of the collection  $\mathcal{T}$  which is induced by phenomenal unity. Furthermore, they appear to satisfy the axioms of a topology as stated above: the whole set of positions in a visual experience is a region; it seems to be the case that intersections of regions in visual experience are also regions in visual experience; and it seems to be the case that the union of any two regions in visual experience is also a region in visual experience. For the empty set, no  $S$ -aspect of consciousness is required (there are no relations of the corresponding unary relation), so we may take the empty set to be a member of  $\mathcal{T}$ . Thus, all axioms of a topology are satisfied.

Therefore, if we take  $M$  to denote the position aspects of visual experiences, and choose  $\mathcal{T}$  to comprise the regions of visual experience, then  $(M, \mathcal{T})$  is indeed a topological structure. And, as shown above, it is a structure of conscious experience as defined in (MSC). We thus find that, because of the subsumptive unity thesis, this topological structure is indeed a mathematical structure of conscious experience; much like conjectured in Tallon-Baudry (2022), it is a topology of the visual content of subjective experience.

## 5 The three problems revisited

In this section, we discuss how the new approach (MSC), which we have developed in Sect. 2.2, resolves the three problems discovered in Sect. 1.

### Problem 1: Incompatible structures

The first problem was that the condition (MDC), which has been applied in previous approaches, admits incompatible structures to conscious experience. Is this also true of (MSC)?

If two structures are incompatible, then there exists at least one automorphism of one structure that is not an automorphism of the other structure.<sup>14</sup> As we explain below, this condition implies that two incompatible structures cannot have all  $S$ -aspects in common. Therefore, it is not possible for two incompatible structures to pertain to conscious experience in the exact same way; so, (MSC) indeed resolves the problem of incompatible structures.

Let  $S$  and  $S'$  denote two incompatible structures (in the narrow sense of the term) with the same domains. Then, there is at least one automorphism of one structure that is not an automorphism of the other structure. Let us denote such an automorphism

<sup>13</sup> It is also plausible to think that visual experiences do not contain positions as aspects, but only regions. However, assessing whether or not this is the case goes beyond the scope of this paper. Here, we assume that positions are aspects of visual experiences.

<sup>14</sup> Automorphisms are structure-preserving mappings from a structure to itself. Put in terms of the terminology we have introduced in Sect. 2.2, automorphisms are mappings  $v$  that map the domains of a structure to themselves. These mappings have to be bijective, and they have to preserve the structure, meaning that they have to satisfy (P1) for all elements of the domain in case of relations, and (P2) for elements of the domains in the case of functions.

by  $v$  and assume that it is an automorphism of  $S$  but not of  $S'$ . Because  $v$  is not an automorphism of  $S'$ , it follows that there is at least one set of relata  $b_1, \dots, b_m$  of  $S'$  in some  $A(e)$ , such that the variation  $v : A(e) \rightarrow A(e)$  induced by the automorphism does not preserve  $S'$  with respect to these relata. On the other hand, because  $v$  is an automorphism of  $S$ , it follows that this variation preserves  $S$  with respect to  $b_1, \dots, b_m$ . If an aspect  $a$  is an  $S'$ -aspect, then, applying the definition of  $S'$ -aspects, we find that the variation  $v$  needs to change it. In contrast, if an aspect  $a$  is an  $S$ -aspect, then, applying the definition of  $S$ -aspects, we find that the variation  $v$  must not change it; either because the  $b_1, \dots, b_m$  do not constitute relata of  $S$ , or because the variation  $v$  preserves  $S$  with respect to relata  $b_1, \dots, b_m$ . Because an aspect cannot be both changed and not changed under a single variation, there cannot be an aspect  $a$  that is both an  $S$ -aspect and an  $S'$ -aspect.

### Problem 2: Arbitrary re-definitions

The definition (MSC) also resolves the problem of arbitrary re-definitions. That's because any re-definition changes the relations or functions of the respective structure, and therefore generates an own, independent condition for something to be an  $S$ -aspect of the redefined structure. Whether or not this new  $S$ -aspect is a part of conscious experience is a substantive question that depends on the actual experiences of the subject under consideration; it is not automatically the case.

Consider, as examples, the cases of rescaling a metric, which we have introduced in Sect. 1. If, per assumption,  $(M, d)$  were a structure of conscious experience, then for any relata  $(b_1, b_2, d(b_1, b_2))$ , the condition for  $d$ -aspects would have to be satisfied. Rescaling this to  $(M, C \cdot d)$  generates a new condition because now, the relata to be considered are  $(b_1, b_2, C \cdot d(b_1, b_2))$ . These are different relata, and correspondingly, different experiences and different variations will enter the definition of a  $C \cdot d$ -aspect. The same is true for an  $(f(a) + f(b)) \cdot d(a, b)$ -aspect. Whether or not these structures satisfy (MSC) depends on the details of the conscious experiences under consideration; but they do not automatically satisfy (MSC) just because  $(M, d)$  does.

### Problem 3: Indifference to consciousness

The third problem is resolved, finally, because of the introduction of  $S$ -aspects in (MSC), which are a counterpart “in” conscious experience to the structure in the narrow sense of the term.  $S$ -aspects introduce a connection between functions or relations in a mathematical structure, on the one hand, and aspects (qualia, qualities, or phenomenal properties) of conscious experiences, on the other hand. Because  $S$ -aspects are part of the definition of (MSC), any application of (MSC) requires engaging with details of the conscious experiences of the subject under consideration; (MSC) is not indifferent to conscious experience in the sense of Problem 3 of Sect. 1.

Consider, for example, the two topological structures of Sect. 1. While (MDC) only required us to check whether the structures address aspects and satisfy the axioms, (MSC) also requires us to check whether there is an  $S$ -aspect in conscious experience that corresponds to the topological structures. As we have seen in Sect. 4, this involves

a careful investigation of conscious experience and relies on intricate notions such as phenomenal unity.

## 6 Conclusion

In this article, we investigated mathematical structures and mathematical spaces of conscious experience. We were not concerned with questions of type or explicit form of these structures or spaces, but with the question of what it means to speak about mathematical structures or mathematical spaces of conscious experiences in the first place. We answer this question by providing a definition of what *mathematical structures of conscious experience* are. This definition provides a foundation for the construction, investigation and identification of concepts like phenomenal spaces, quality spaces, qualia spaces and  $Q$ -structures.

Our definition of mathematical structures of conscious experiences is grounded in a foundational understanding of mathematical structures and spaces as laid out by mathematical logic. And it is axiomatic in the sense that it can be applied to any conceptualization of conscious experiences, and any choice of aspects thereof (e.g. qualia, qualities, phenomenal properties, phenomenal distinctions), which satisfy the formal requirement that for every conscious experience there is a well-defined set of aspects.

Our definition rests on the notion of *variations*, which are changes of one conscious experience to another. Because variations can be induced introspectively (for example, as in Husserl's imaginary variations), stimulated in a laboratory by change of stimuli, or studied theoretically based on a proposed theory of consciousness, our definition constitutes a general method to identify and study structures of conscious experience.

The grounding of mathematical structures of conscious experiences proposed here is *methodologically neutral* in the sense that it can be combined with many methods, practices, and procedures that are used to investigate conscious experience, spanning empirical, analytical, and phenomenological research. Furthermore, it is *conceptually neutral* in the sense that it can be applied to any conception of 'conscious experience' and 'aspects' thereof, as long as every conscious experience comes with a well-defined set of aspects. This includes common conceptions using qualities, qualia, or phenomenal properties, but also less common ideas based on atomistic conceptions of states of consciousness or phenomenal distinctions.

Our definition complements recent approaches that study quality spaces, qualia spaces, or phenomenal spaces, because it retains the abstract condition that these proposals apply—Condition (MDC) in Sect. 1—as a necessary part. This abstract condition is extended by our proposal, so as to avoid three problems that interfere with recent approaches, see Sect. 1.

In light of the increasing interest in using mathematical structures to model and represent conscious experiences in the scientific study of consciousness and philosophy of mind, the investigation of how to define and understand mathematical structures of conscious experience is important, in our view. This work contributes to this investigation. It highlights issues with previous ways of understanding structural claims and offers an improved conception that rests on meaningful desiderata. Hence, we

hope, it contributes to building a foundation for structural research for both theory and experimental practice.

As a first application, and to illustrate our definition, we considered *relative similarity* and *topological spaces*. We found that relative similarity, which plays an important role in several constructions of quality spaces, is indeed a mathematical structure of conscious experience, see Sect. 3. Topological spaces also qualify as mathematical structures of conscious experience, but for a surprising reason: they are intimately related to phenomenal unity, see Sect. 4.

We view the results presented here as one further step in a long journey to investigate conscious experience mathematically. This step raises new questions and creates new opportunities, both of which can only be explored in an interdisciplinary manner. A new question, for example, is whether our result on mathematical structures might open new perspectives on measurements of consciousness (Irvine, 2013), as arguably promised by the Representational Theory of Measurement (Krantz et al., 1971) whenever an axiomatic structure on a target domain is available. We hope that, ultimately, our result provides a basis for developing a common formal language to study consciousness across domains.

**Acknowledgements** We would like to thank the participants of the 2022 *Modelling Consciousness Workshop* and of the *Models of Consciousness 3* conference, both organized under the umbrella of the Association for Mathematical Consciousness Science, as well as members of the *Munich Center for Mathematical Philosophy* for fruitful discussions and helpful comments, and in particular Jonathan Mason, Robin Lorenz, Ian Durham, Christian List, Andrew Lee, Etienne Jacques and Pedro Resende for valuable feedback on the manuscript. This research was supported by grant number FQXi-RFP-CPW-2018 from the Foundational Questions Institute and Fetzer Franklin Fund, a donor advised fund of the Silicon Valley Community Foundation. We would like to thank the Dutch Research Council (NWO) for (partly) financing TL's work on project number 182.069 of the research programme Fluid Spintronics, and the Mathematical Institute of the University of Oxford for hosting JK while working on this project.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

## Declarations

**Conflict of interest** The authors declare that there are no financial or non-financial competing interests that are directly or indirectly related to the work submitted for publication.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Bayne, T. (2005). Divided brains and unified phenomenology: A review essay on Michael Tye's consciousness and persons. *Philosophical Psychology*, 18(4), 495–512.
- Bayne, T. (2012). *The unity of consciousness*. Oxford University Press.

- Bayne, T. J., & Chalmers, D. J. (2003). What is the unity of consciousness? In A. Cleeremans (Ed.), *The unity of consciousness*. Oxford University Press.
- Chalmers, D. J., & McQueen, K. J. (in press). 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.
- Clark, A. (2000). *A theory of sentience*. Clarendon Press.
- Cleeremans, A., & Frith, C. (2003). *The unity of consciousness*. Oxford University Press.
- Coninx, S. (2022). A multidimensional phenomenal space for pain: Structure, primitiveness, and utility. *Phenomenology and the Cognitive Sciences*, 21(1), 223–243.
- Fechner, G. (1860). *Elements of psychophysics* (Vol. I). Holt, Rinehart and Winston.
- Fink, S. B., Kob, L., & Lyre, H. (2021). A structural constraint on neural correlates of consciousness. *Philosophy and the Mind Sciences*, 2. <https://doi.org/10.33735/phimisci.2021.79>
- Fortier-Davy, M., & Millièrè, R. (2020). The multi-dimensional approach to drug-induced states: A commentary on Bayne and Carter's "dimensions of consciousness and the psychedelic state". *Neuroscience of Consciousness*, 2020(1), niaa004.
- Gert, J. (2017). Quality spaces: Mental and physical. *Philosophical Psychology*, 30(5), 525–544.
- Grindrod, P. (2018). On human consciousness: A mathematical perspective. *Network Neuroscience*, 2(1), 23–40.
- Haun, A., & Tononi, G. (2019). Why does space feel the way it does? Towards a principled account of spatial experience. *Entropy*, 21(12), 1160.
- Hoffman, D. D., & Prakash, C. (2014). Objects of consciousness. *Frontiers in Psychology*, 5, 577.
- Husserl, E. (1936/1970). *The crisis of European sciences and transcendental phenomenology: An introduction to phenomenological philosophy*. Northwestern University Press.
- Irvine, E. (2013). Measures of consciousness. *Philosophy Compass*, 8(3), 285–297.
- Joshi, K. (1983). *Introduction to general topology*. Wiley Eastern.
- Joshi, K. D. (1989). *Foundations of discrete mathematics*. New Age International.
- Kleiner, J. (2020a). Brain states matter. A reply to the unfolding argument. *Consciousness and Cognition*, 85, 102981.
- Kleiner, J. (2020b). Mathematical models of consciousness. *Entropy*, 22(6), 609.
- Kleiner, J., & Hoel, E. (2021). Falsification and consciousness. *Neuroscience of Consciousness*, 2021(1), niab001.
- Kleiner, J., & Tull, S. (2021). The mathematical structure of integrated information theory. *Frontiers in Applied Mathematics and Statistics*, 6, 74.
- Klincewicz, M. (2011). Quality space model of temporal perception. In *Multidisciplinary aspects of time and time perception* (pp. 230–245). Springer.
- Kostic, D. (2012). The vagueness constraint and the quality space for pain. *Philosophical Psychology*, 25(6), 929–939.
- Krantz, D., Luce, D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement* (Vol. I–III). Academic Press.
- Kuehni, R. G., & Schwarz, A. (2008). *Color ordered: A survey of color systems from antiquity to the present*. Oxford University Press.
- Lee, A. Y. (2021). Modeling mental qualities. *Philosophical Review*, 130(2), 263–298.
- Lee, A. Y. (2022). Objective phenomenology. *Erkenntnis*, 1–20. <https://doi.org/10.1007/s10670-022-00576-0>
- Lee, A. Y. (2023). Degrees of consciousness. *Noûs*. *Noûs*, 57, 553–575. <https://doi.org/10.1111/nous.12421>
- Lyre, H. (2022). Neurophenomenal structuralism. A philosophical agenda for a structuralist neuroscience of consciousness. *Neuroscience of Consciousness*, 2022(1), niac012.
- Mason, J. W. (2013). Consciousness and the structuring property of typical data. *Complexity*, 18(3), 28–37.
- Mason, J. W. (2021). Model unity and the unity of consciousness: Developments in expected float entropy minimisation. *Entropy*, 23(11), 1444.
- Mileti, J. (2022). *Modern mathematical logic*. Cambridge University Press.
- 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.
- Prentner, R. (2019). Consciousness and topologically structured phenomenal spaces. *Consciousness and Cognition*, 70, 25–38.
- Renero, A. (2014). Consciousness and mental qualities for auditory sensations. *Journal of Consciousness Studies*, 21(9–10), 179–204.

- Resende, P. (2022). Qualia as physical measurements: A mathematical model of qualia and pure concepts. arXiv preprint [arXiv:2203.10602](https://arxiv.org/abs/2203.10602)
- Roelofs, L. (2016). The unity of consciousness, within subjects and between subjects. *Philosophical Studies*, 173(12), 3199–3221.
- 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.
- Rosenthal, D. M. (2016). Quality spaces, relocation, and grain. In O'Shea (Ed.), *Sellars and his legacy* (pp. 149–185). Oxford University Press.
- Rudin, W. (1976). *Principles of mathematical analysis* (Vol. 3). McGraw-Hill.
- 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*, 26(12), 1068–1069.
- 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.
- Tsuchiya, N., & Saigo, H. (2021). A relational approach to consciousness: Categories of level and contents of consciousness. *Neuroscience of Consciousness*, 2021(2), niab034.
- Tye, M. (2021). Qualia. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2021 ed.). Stanford University. <https://plato.stanford.edu/archives/fall2021/entries/qualia/>
- Wiese, W. (2018). What is it like to experience a third man? The phenomenological Bradley and how to solve it. In W. Wiese (Ed.), *Experienced wholeness: Integrating insights from Gestalt theory, cognitive neuroscience, and predictive processing*. The MIT Press. <https://doi.org/10.7551/mitpress/9780262036993.003.0004>
- Yoshimi, J. (2007). Mathematizing phenomenology. *Phenomenology and the Cognitive Sciences*, 6(3), 271–291.
- Young, B. D., Keller, A., & Rosenthal, D. (2014). Quality-space theory in olfaction. *Frontiers in Psychology*, 5, 1.
- 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.