Adapting fundamental concepts from insect societies to the demands of bio-ontologies.

Thiago S. R. Silva[1,3]* & Marcio R. Pie[2,3]

[1] School of Biological Sciences, The University of Hong Kong. Hong Kong SAR, China.

[2] Biology Department, Edge Hill University. Ormskirk, United Kingdom.

[3] Departamento de Zoologia, Universidade Federal do Paraná. Curitiba, Brazil.

*Corresponding Author's Email: tsranzanidasilva@gmail.com

Abstract

The present work provides a logical account for translating fundamental concepts in the study of social insects to bio-ontologies, namely 'colony' and 'caste', which can be used by researchers for describing and organizing entities belonging to non-human societies, as well as for the provision of evidential criteria for evaluating cumulative-constitutive explanations of 'social' entities. We explore how different properties can provide connectedness of 'colonies' and 'castes' and discuss the implications of this approach to biological representation and ontology design. Finally, we briefly discuss some limitations in current approaches to demarcate subcategories of 'castes'.

Keywords

Biological entities, frames of reference, causal unity, realist ontology, knowledge systems.

Introduction

Terminological standardization has been a matter of debate among research groups that study insect societies for several years [1]–[6]. According to several authors (e.g. [2], [6]–[8]), there are distinct reasons for the need for terminological univocity in this field of inquiry: (i) to provide a clearer conceptual landscape that enables researchers to explore drivers in the evolution of sociality in animals, (ii) to provide the linguistic components for a more straightforward communication, whether textually or orally, and (iii) to provide a framework for unified concept representation and organization, enhancing data comparability and reproducibility. Aiming for concept clarity is an important first step for standardization of specialized terminologies, although it does not directly entail standardization *per se*, since other terminological dimensions must be explored for consolidating shared sets of terms [9]–[11]. While the above-mentioned

authors (and many others) provided distinct accounts on defining criteria for establishing unambiguous concepts of social categories in insect societies, few of them have tried to explore how these categories can be translated into data models. While urging for 'holistic' approaches for defining and describing social categories, researchers generally opt for combined frames of reference when defining them, namely spatiostructural and functional frames [5], [8], [12], [13].

Although employing different frames of reference during concept representation can be considered the best approach for demarcating biological entities, researchers can equivocally ascribe concepts to certain frames in this process, leading to category mistakes [14]. A category mistake is a semantic or ontological error made by an enunciator when it represents the facts of mental life as if they belonged to one logical entity or category (or domains of entities) when they actually belong to another [14]. Examples of category mistakes in 'caste' representation include gamergates and ergatoids, as commonly understood in myrmecology [15]–[17] (Figure 1). Traditionally, 'castes' in ants have been defined by a combination of properties belonging to two distinct frames of reference (i.e., spatiostructural and functional); workers are sterile apterous females (Figure 1.A) and queens are fertile macropterous females (Figure 1.B) [16]–[20]. In this particular example, gamergates would be subsumed in workers while ergatoids are subsumed in queens; additional defining properties, however, are conflicting with the properties of their supervening categories (Figure 1.C and D).

Figure 1. Example of category mistake in two 'caste' concepts applied in myrmecology: gamergate and ergatoid. Equivocal relations leading to category mistakes are marked as dotted lines. In this example, most relational properties are not necessarily compliant with formal ontologies (e.g., *incapable of*, *determined by presence of*, *has anatomic quality*) and are used only for didactic purposes.

According to [21], researchers are partly led to construe mental concepts as signifying occurrences of underlying processes because they conflate how they explain an individual's successful moves with what they require of the individual in making those moves. Although that can be certainly true, another possible explanation for this phenomenon is that researchers do not necessarily are aware of the frames of reference used to elaborate concepts. Even though the causes for category mistakes are not always so easily accounted for [22], it is important to consider its implications in the life sciences, especially if we intend to explore compositional approaches when representing biological concepts. Applying a logical framework that incorporates distinct frames of

reference, relations of causal unity and levels of organization can help researchers account for category mistakes, at least to some degree.

The aim of this work is to provide a logical account for translating two fundamental concepts in the study of social insects to a realist ontology – i.e., 'colony' and 'caste' –, which can be used by researchers for describing and organizing entities belonging to non-human societies, as well as for the provision of evidential criteria for evaluating constitutive explanations of 'social' entities. We will argue that both concepts can be demarcated under the three most common frames of reference used in biological sciences, according to [23]: spatiostructural, functional, and historical/evolutionary. We will adjust and refine previous propositions for a 'caste' ontology while addressing some misconceptions made by [5]. Initially, we will briefly address some fundamental properties used for demarcating 'colonies' of social insects as natural units. We will then discuss the fundamental properties that can be used to demarcate 'castes' and how they contribute to the overall connectedness of 'colonies'. Lastly, we will provide a brief account on limitations of current approaches to demarcate subcategories of 'castes'.

Methods

Specialized terms

In this work, we used quotation marks when referring to the names 'colony' and 'caste', in order to highlight the inappropriateness of its usage for communication in contemporary biology (cf. discussion at [24]–[26]). We understand that reevaluating the linguistic component of terminological units – i.e., multidimensional units composed of linguistic, cognitive, and situational components [27] – is essential for the development of a more inclusive and diverse science. We suggest that a reevaluation of this magnitude should be conducted by a diverse group of researchers through a transdisciplinary approach, which is a desirable condition needed to address linguistic modifications. Since it is not the aim of the present work to reevaluate the name usage throughout the history of the discipline, we refrained from providing alternative terms to the concepts that refer to aggregations of individual organisms belonging to groups of social insects and to distinct categories of the phenotypes in non-human societies, lest we change an oppressive-laden name by another.

Frames of reference and relations of causal unity

We used realist-based approaches to address the possibilities of adapting fundamental concepts used in the study of social insects for application in bio-ontologies. The idea is to create extensions of already-existing entities that are conformant to the Basic Formal Ontology (henceforth BFO). To achieve this, we had to deal with two basic problems regarding natural kinds: (1) granularity dependence of bona fide boundaries and (2) perspective dependence of bona fide boundaries. Both problems relate to discontinuity (and, conversely, connectedness) and qualitative heterogeneity, at least to a certain degree [28]. To deal with these issues, we demarcated 'colonies' and 'castes' by their specific types of causal unities, which, in turn, are respectively connected to specific frames of reference [23], [29], [30].

To specifically address the mind-independence of non-structural boundaries of 'colonies' and 'castes', we followed the examples of functional and historical/evolutionary units provided by [31], which can be divided in (i) dispositions independent of morphogenesis and (ii) dispositions involving morphogenesis for the first type, and (iii) common causal history (*i.e.* structural integrity and stability of an entity over time) and (iv) common historical origin (*i.e.* constituent historical relations of entities distributed in time and space) for the second type.

Results

Results 1: 'Colonies' of social insects as unified causal systems

Insect 'colonies' are considered as aggregates of single-species organisms that occupy a given geographical space and share some degree of kinship, meaning that its members normally have high historical (i.e., genealogical) relatedness among themselves and live close to each other [32]. Every individual belonging to a 'colony' is a member part of this aggregate and, thus, they are considered as its proximal parts, which are 'connected' to one another by some sort of 'natural bond' (in our case, presented in the form of relations of causal unity). This means that every termite individual is a member part of a termite 'colony', for example, and that this membership relation is what determines the 'colony' as an aggregation of organisms.

There have been some studies that sought to establish the underlying properties of social insect 'colonies' under the 'complex adaptive system' (CAS) paradigm [33], [34] – i.e., systems were properties in coarser levels of organization emerge from interactions among components belonging to finer levels of organization and may feedback to influence the subsequent development of those interactions [35]. The use of

CAS is very promising to partially explain ‘colonial’ connectedness through time and space, especially when theories of social organization are used as an approach to explore colonial dynamics. There are some limitations, however, in the idea of ‘colonies’ as CAS, because they are normally explored as means to explain emergence merely through dispositional properties (i.e., emergence of spatiotemporal structures, like trails, towers, nests, bivouacs, through non-morphogenetic dispositions; e.g., [34], [36]–[38]), while emergence explanations of spatiostructural and historical/evolutionary properties are non-compliant with the model.

To address these limitations, we work under the premise that ‘colonies’ are causally relatively isolated entities (i.e., they are unified causal systems isolated from their surroundings [39], [40]) meaning that they are both structured through and maximal relative to certain types of causal unities. In this sense, insect ‘colonies’ would be considered as natural units, existing independent of human partitioning activities. Ideally, there are five types of causal unities that are believed to promote connectedness of a given natural unit: (i) causal unity via physical covering; (ii) causal unity via internal physical forces; (iii) causal unity via engineered assembly of components; (iv) causal unity via bearing a specific function; and (v) causal unity via common historical/evolutionary origin [23], [30]. It is worth mentioning that the first three types of causal unities are not ontologically independent from one another, where (i) and (iii) are existentially dependent and thus supervene on (ii) [23].

Ideally, ‘colonies’ of social insects are maintained as spatiotemporal connected units thanks to some properties that allow them to achieve some degree of internal homeostasis [41]. The main properties that enable ‘colonial’ connectedness are chemical space, self-organization, and resilience. Each one of these properties can be linked to a specific frame of reference and the parts that enact them can be further partitioned in their respective perspectives. In the following sections, we will show how chemical space, self-organization, and resilience can be connected to spatiostructural, functional, and historical/evolutionary frames of reference, respectively.

Results 1.1. Spatiostructural frame of reference and structural properties of ‘colonies’

Historically, structural properties of ‘colonies’ of social insects have been directly linked to the properties of the structures that they occupy, namely nests [42]. However, the structural properties of ‘colonies’ differ substantially from the structural properties

of the physical niches they inhabit, although partly depending on them to maintain their connectedness.

In the case of insect ‘colonies’, nests and ‘living structures’ would not necessarily be considered as physical coverings in the sense of [30], since they are not comprised solely of an enclosing physical structure, but rather are considered as niche tokens [43] in the form of double hole structures [44]; for holes see [45]. In this sense, both nests and ‘living structures’ are composed of a central hole, which represents the ‘colony’ proper (also called the tenant) and their accompanied medium (i.e. the displaced air that is precisely the right size and shape to be occupied by its individual members), a surrounding medium (or environing hole), which represents the space used by the ‘colony’ to move, breathe, and perform activities (i.e. chambers, shafts, tunnels, floors and/or trails), and, in most cases, an external enclosing structure (or retainer), which surrounds both the central hole and the environing hole.

The surrounding medium is an important component that allows for the connectedness of the ‘colony’, given that it allows for the occurrence of several causally-linked biological processes. For instance, the medium provides a space for thermoregulation, regardless of the presence of an abiotic physical retainer. As an example, in army ants, bivouac temperature at its core is maintained stable and higher than the environment temperature [38], through higher respiratory rates of workers when they are gathered [46]; this property allows for a fast development of the immature stages and an overall stability of the ‘colony’ in the long-run. In this particular case, the crevices that compose the core of the ‘colony’ (i.e., the spaces between each individual organism) are filled by a mixed medium (i.e., air plus the metabolic products of respiratory processes) which provides the ideal thermal conditions for maintaining their respective tenants.

More importantly, however, is that the mixed medium in insect ‘colonies’ allows for a physical connection among its parts in the form of weak physical bonds at the molecular level. In this case, the mixed medium is filled with volatile components, namely pheromones, which are responsible for physically connecting each member of the ‘colony’ to its whole. Volatile chemical components and/or chemical components attached to an abiotic surface (such as trails outside the nest or in tunnels within the nest) provide a form of connectedness between the members of the ‘colony’, since pheromone production and recognition involves a series of physical-chemical causal relations occurring in succession, leading to the connection of two or more organisms of

the aggregate. The process of pheromone production and recognition is considered to provide physical connection because each component involved in the process is structurally bound (normally in the form of weak interactions) to the immediately next component in the chain of events that corresponds to the production-recognition process.

There are a number of studies discussing the importance of chemical distance and its relation with increase in aggression among congeneric species of social insects, but also highlighting the importance of variation of cuticular hydrocarbons that are 'caste' and sex specific [47]. Chemical signatures are a major force promoting physical connectedness of insect aggregations, independent of their dispositions; specific patterns of chemical structure (i.e., chemical profiles) are recognized by the members of a given colony and these members decide how they must respond when faced with this variation (whether in an aggressive manner or not). If the pattern of the 'colonial' chemical structure is dissimilar enough, or if its concentration is lower than a given threshold, disconnection unfolds through aggressive behaviors or exclusion of the member from the colony (see Results 2.1 for a more comprehensive account on the role of chemical profiles among different 'castes').

There are particular cases in which the 'colony' can be considered to be fragmented, given that all the members of the aggregate do not necessarily share the same spatial environment (i.e., do not share the same retainer-delimited space). In polydomous ant 'colonies' (i.e., same-species aggregates of social insects that inhabit more-than-one unconnected nests), some members of a 'colony' are spatially separated from other members by disconnected nests. Nonetheless, they share the same chemical space.

Since the surrounding medium allows for a sort of physical connection between the members of a 'colony', in the form of chemical bonds acting in a causal flow, the 'colony' is considered to be structured through and maximal relative to a relation of causal unity via internal physical forces (Figure 2.A; Table 1). In this case, the internal physical forces are considered almost literally as presented in the examples by [30], as in the '(…) *fundamental forces of strong and weak interactions, by covalent or ionic bonds or more generally by forces of a type which makes the overall sum of forces strong enough to act in such a way as to maintain the group together relative to the strength of attractive or destructive forces in its environmental neighborhood*'.

Figure 2 Schematic representation of different types of properties that allow for 'colonial' connectedness and their respective frames of reference. A. 'Colonial' connectedness through causal unity via internal

physical forces (i.e., chemical space); B. 'Colonial' connectedness through causal unity via physiological dispositions (i.e., ecological interactions); C. 'Colonial' connectedness through causal unity via historical origin of its parts (i.e., relations of common kinship).

Table 1. Ontological criteria for demarcating fundamental concepts in the study of social insects to the demands of bio-ontologies, following a realist framework.

	'COLONY'	**'CASTE'**
SPATIOSTRUCTURAL PROPERTY	'Colonial' chemical profiles	'Caste' chemical profile?
FUNCTIONAL PROPERTY	Physiological disposition	Developmental disposition
HISTORICAL/EVOLUTIONARY PROPERTY	Relation of common kinship	Developmental history

It is important to consider, however, that insect 'colonies' are dynamic systems that vary on their physical connectedness through time. This is especially true when a single individual (or a group of individuals that are a subset of the 'colony') lose their physical connectedness with the whole during a given period of time – i.e., when the individual leaves the pheromone space of the colony. This limitation will be accounted for when we discuss resilience and how it promotes 'colony' connectedness (see Results 1.3).

Results 1.2 Functional frame of reference and functional properties of 'colonies'

Considering 'colony' regulation promoted by self-organization, dispositional properties promote spatial restructuring and connectedness of 'colonies' when faced with local environmental/ecological constraints. Self-organization is an important attribute for insect societies, since it allows for the development of the aggregate, improving its overall fitness. Both spatial organization (in the form of movement and density) and social organization (in the form of task allocation) are fundamental for the regulation of a 'colony'. Spatial organization is believed to be an important mechanism for colonial regulation, since it allows for the aggregate to circumvent pitfalls of excessive growth during extremely favorable environmental conditions, or to avoid colonial collapse when faced with suboptimal or catastrophic situations. In this sense, spatial organization is intrinsically related to social organization, since it is believed that interaction rates are typically density-dependent, meaning that individual movement, 'colony' size and 'colony' density significantly alter group dynamics [48].

One interesting example of colonial self-organization properties can be found in trail formation in ants. When a food resource is available (e.g., a piece of cake left

unattended on the kitchen's table) inevitably a forager will discover it, especially if the resource is at the vicinities of the ant nest. Shortly after discovering the piece of cake, the forager will recruit other ants; the number of ants will increase exponentially until there are enough foragers needed to efficiently obtain the food resource. This regulation of individuals during trail formation is understood as a system comprised of both positive and negative feedbacks: new arrivals at the food source are considered as 'births', while abandonments of food sources are considered as 'deaths'. The recruitment ceases when the system reaches a 'stable population size', meaning that the 'birth rate' equals to the 'death rate'. These negative/positive feedback loops are considered important dispositions of self-organization in social insects. Given that the ants involved in the trail formation interact with the biotic and abiotic environment, the 'colony' disposition, in this case, is independent of morphogenesis.

According to [49], members of a 'colony' are known to exhibit several distinct movement zones, which have been linked to an individual's behavioral repertoire in a variety of social insects [50]–[56]. These movement zones (spatial fidelity zones according to [57] are thought to play a key role in labor division, since they promote task specialization and reduce task switching costs. Therefore, both social and spatial organization play an important role in maintaining 'colonial' stability when 'colonial' demography fluctuates or environmental pressures constraint 'colonial' connectedness [49]. Hence, changes in disposition of members of the 'colony' (in the form of interactions) are responsible for stabilizing it. Therefore, 'colonies' are structured through and are maximal relative to physiological dispositions, acting as functional units of ecology (Figure 2.B; Table 1).

Results 1.3 Historical/evolutionary frame of reference and historical/evolutionary properties of 'colonies'

Considering 'colony' regulation promoted by self-recognition and kin-recognition, genealogical properties promote integrity and connectedness of 'colonies'. During colonial activities, individuals recognize themselves and other members of the 'colony' through a series of visual and/or olfactory cues that indicate they share the same genetic pool [32]. Through kin-recognition conflict is mostly thoroughly avoided, favoring cooperative behavior. Hence, relations of common kinship bound groups of organisms across their spatiotemporal boundaries, meaning that these relations provide a sort of

connection to spatiotemporally scattered groups of organisms in the form of genealogical lineages.

'Colony' demography, which is an important property for 'colony' stability, is also promoted by self-recognition and kin-recognition. Depending on certain intrinsic and extrinsic situations (i.e., size and composition of 'colony', and environmental factors, respectively), members of the 'colony' modulate the production or expression of certain phenotypes. The expression of certain phenotypes provides the means for the 'colony' to growth or to reestablish itself when faced with some external perturbation. Hence, 'colony' stability and connectedness are intrinsically related to self-recognition and kin-recognition promoted by common kinship and, thus, 'colonies' are structured through and maximal relative to the historical origin of its parts through relations of common kinship (Figure 2.C; Table 1).

Results 2: 'Castes' of social insects as unified causal systems

Considering the accounts on 'colonial' connectedness presented in the previous sections, we are led to consider 'castes' as properties of a given 'colony', in the form of dispositions to produce individuals of certain phenotypes as response to intrinsic and/or extrinsic stimuli. From that, we can assume that 'castes' are results of the 'colony's gene pool when they react to certain trigger factors, changing the structure of the distribution pattern of different phenotypes in it. 'Castes' then are determined according to the 'colony' internal demands when constrained by both external and internal forces.

Taking into consideration the demands for 'colonial' connectedness mentioned in the previous sections, it is possible to ascertain that there are several drivers for 'caste' determination. As properties of a 'colony' that exist to provide internal stability for this unit, 'castes' also can be considered as natural units themselves. Hence, they exist independent of human partitioning activities, similar to other biological entities, being causally relatively isolated and both structured through and maximal relative to certain types of causal unities.

Contrary to 'colonies', 'castes' are determined at less inclusive levels of biological organization (organismal and lower levels) rather than more inclusive levels (such as in aggregations of organisms). Nonetheless, the result of the determination process is an individual possessing certain properties that can contribute directly or indirectly to the connectedness of the 'colony'. Equally important, however, is how these properties contribute to the connectedness of the 'caste' itself.

Results 2.1 Spatiostructural frame of reference and structural properties of 'castes'

Although 'castes' can be demarcated through a spatiostructural frame of reference, it is not straightforward procedure to define a relation of causal unity which maintains their spatiostructural connectedness. Normally, we would establish the spatiostructural connectedness of 'castes' using the same type of causal relations that are used for organisms (i.e., cumulative-constitutive relations of internal physical forces of the structures that compose the organism, in the different levels of organization) and, as an operational criterion, we would use chemical profiles to demarcate 'castes'. It is uncertain if chemical profiles promote spatiostructural connectedness in 'castes' like it does for 'colonies' (Table 1). It could be argued that, as byproducts of all the developmental processes that culminate in 'caste' determination, chemical profiles would provide the causal element for spatiostructural connectedness. However, insofar, 'caste'-specific chemical profiles had been known to allow kin-recognition and as a mediating component for intra-colonial interactions exclusively, which would contribute to connectedness at the 'colonial' level (see Results 1.2 and 1.3).

Although they are considered to promote connectedness at the 'caste' level, chemical profiles can be considered as a useful operational criterion to demarcate 'caste' categories. As mentioned before, since they are byproducts of 'caste' determination, there are some pheromones that are considered to be 'caste'-specific and, as such, can be used to differentiate the distinct 'caste' categories.

Results 2.2 Functional frame of reference and functional properties of 'castes'

'Castes' are determined by several developmental changes induced by both genetically and environmentally constraints, transforming its spatiostructurally make-up at the end of the process. Since 'caste' determination depends on several genetically determined parts, and those parts are intrinsically dependent of one another (i.e., often the development of one part is causally dependent on the development of another part or both developments are controlled by the same genetic control mechanism), their development are causally linked and coordinated with one another, making 'castes' functional units of development. More specifically, the determination of certain phenotypic traits is influenced by several processes acting in finer levels of biological organization. Said traits can be determined equally or individually by epigenetic

processes [58], physiological determinants [59], and/or environmental constraints [60]. Hence, each component on a specific level of organization has a differential participation on the determination of particular 'castes'. Those components can (a) individually and directly influence the determination of the 'caste', or (b) synergistic operate along with other components of the same level of organization they belong to, composing a coarser level entity that, in turn, will directly influence the development of the 'caste'.

For example, one species of desert ant, *Cataglyphis mauritanica* (Emery, 1906a), has a strong genetic bias to 'caste' development, such that, under normal circumstances, certain genotypes always develop in small workers [61]. However, juvenile hormone treatment causes worker-destined genotypes to develop into larger queens, while queen-destined genotypes will develop into small workers when reared in small 'colonies', where the larvae are likely starved [62]. Hence, the determination of 'castes' can be directed on the genotype level in some cases, while being directed at the metabolic or environmental level in other cases.

Therefore, 'castes' are defined as functional units of development because they have differential dispositions for determination. It is important to highlight that genetic dispositions for 'caste' determination are hardly fixed in most social insects. Most of the time, discrete phenotypes arise from a single genotype because of different environmental conditions. Each 'caste' would correspond to a functional unit of development (Table 1), i.e., workers correspond to one functional unit while soldiers correspond to another functional unit, because they can be independently characterized by the different morphogenetic dispositions they bear.

Results 2.3 Historical/evolutionary frame of reference and historical/evolutionary properties of 'castes'

'Castes' are historical units of development since they are composed of particular parts that have proven in the past that they have developed in concert, as a unit of development, maintaining the structural integrity of the entity as a whole – i.e., they exist as a consequence of a series of developmental processes that are causally linked and coordinated to one another (Table 1). Historical units of development differ from functional units of development due to their diachronic property [31]; i.e., while functional units of development are delimited by the propensity of behaving in a certain developmental way in the future, historical units of development are defined by the fact

that they acted as functional units of development in the past. This means that they are based on a common morphogenetic history, opposed to morphogenetic dispositions.

Discussion

Discussion 1: Implications for biological representation and ontology design

Bio-ontologies are important and useful tools not only in the standardization of data and metadata, but also in data integration, data compatibility and comparability, and for data communication and management [63]. However, since several ontologies have been developed with a particular practical purpose in mind [5], with a focus on the definition for very specialized types of entities, definitions for general types of entities are normally lacking. This normally leads to ontological inconsistency and cross-ontology incompatibility. Following a consistent framework for defining biological entities in application and domain reference ontologies is important for enabling cross-ontological compatibility, also enhancing interoperability in future applications.

Since the current proposal is structured in a way that all the biological entities studied need to conform with some fundamental maxims established by BFO, the reevaluation of the ways in which ‘caste’ categories are demarcated is expected. Nonetheless, even if discoveries on ‘caste’ determination are made, this framework enables the reevaluation and evolution of concepts, attribution of new concepts and adjustments of available entities. Considering there are still major knowledge gaps about the drivers of ‘caste’ determination and evolution, along with the huge advances we had in the past few years in fields like eco-evo-devo and thanks to the use of new high-throughput technologies, a versatile yet coherent framework would help us better accommodate new findings, or at least provide enough support for establishing operational criteria to demarcate those entities.

It is also important to highlight that, despite representation of biological entities deriving from cognitive acts – i.e., they are constructed through signs in a representational system – they tend to be correlated with what we understand of ‘biological reality’. In this framework, this is only possible because the biological entities being represented – i.e., ‘colonies’ and ‘castes’ – are demarcated following different frames of reference beyond the spatiostructural one. Additionally, accommodating fundamental concepts in the study of social insects to representation systems such as ontologies allow researchers to focus first on identifying the properties that are believed to provide connectedness to the entities being studied, by means of

relations of causal unity. From there, it should be easier to established operational criteria that would be used to define concepts, which in turn would be used for the provision of explanatory accounts about biological phenomena. This should provide a robust way to avoid certain ontological errors, such as category mistakes.

Lastly, representing ‘colonies’ and ‘castes’ in a more formalized manner is beneficial in the way that we can provide a conceptual framework for debates of downward causation on ‘caste’ expression [64], since it is considered that all individuals in a ‘colony’ share the same genetic background and that ‘caste’ differentiation derives from differences in gene expression [65], [66].

In terms of ontology design, recent efforts bridged important gaps in suggesting additional ontology relations [67], [68]. For example, specifically for ‘castes’, ‘developmentally-related-to’ relations could be used to identify both functional and historical/evolutionary units in their respective frames of reference. However, when modeled this way, we would not necessarily have a universal single-inheritance tree, since each class-subclass relation would be bounded exclusively to their respective frame. This is somewhat problematic, especially if we are concerned with comparability and compatibility of different ontologies. As highlighted by [31], the use of additional types of transitive is-a relations – e.g., ‘is-physiologically-a’, ‘is-developmentally-a’ – could help deal with the issue of organizing different taxonomies within a single-inheritance tree.

<u>Discussion 2: Current limitations for ‘caste’ subcategories</u>

In groups that express intermediate or phenotypic mosaics, or have dispositions to perform specific functions or roles within a ‘colonial’ environment (such as foraging or brood-care), it is common to establish a subcategory of ‘caste’, i.e., ‘subcaste’ [5]. In pragmatic cases – i.e., when researchers have specific investigative motivations for splitting ‘caste’ categories – it would be possible to define some operational criteria, based on the respective frame of reference, to further partition ‘castes’ into ‘subcastes’. However, if we wished to argue that ‘subcastes’ truly represent a subclass of ‘castes’, the former would necessarily need to inherit all properties of the latter and incorporate additional defining properties, thus narrowing in the meaning of the term. However, insofar, attempts to split ‘castes’ into subcategories have not been entirely successful. One prominent example refers to the concept of ‘physical castes’ and ‘temporal castes’ proposed by [69].

According to Wilson's proposition, a 'caste' would be an ensemble of colony members that specialize on particular tasks for prolonged periods of time, being typically (but not necessarily) distinguished by other genetic, anatomical, or physiological traits. He further categorizes 'castes' in two subcategories, namely 'physical castes' and 'temporal castes', with the former being defined by allometric and other anatomic criteria and the latter being defined by age grouping. It is noteworthy to highlight that, although being operationally sound for the most part, Wilson's proposition has some ontological limitations: (i) the absence of a sound framework needed to deal with the issue of connectedness of biological entities in different frames of reference; (ii) ambiguous defining properties of 'castes', which in turn are used to derive the defining properties of the two proposed subcategories; and (iii) arbitrariness in partitioning strategies, leading to creation of several unnested (i.e. not comparable) categories.

The first issue in Wilson's proposition is quite straightforward since connectedness is a pervasive dimension to be considered when demarcating biological entities. Although connectivity is normally understood at the spatiostructural dimension (i.e., one entity physically connected to other neighboring entities), at least in complex systems other types of connection need to be considered. Although Wilson explicitly considers the importance of biological dispositions in defining 'castes' of social insects, he never articulates on how these dispositions promote connectedness of these entities.

Wilson's second issue directly stems from the first one, since identification of defining properties for biological entities depends on a logically oriented framework, with consideration of the distinct frames of reference used to identify said properties. Wilson considers 'castes' as object aggregates (i.e. a group of bona fide epithelially delimited multicellular organisms) that are recognized by a set of functionally defined spatiostructural entities (i.e., anatomic distinct individuals that have the disposition to perform specific biological functions). However, recognition criteria of functionally defined spatiostructural entities cannot be directly inferred from their defining properties (i.e., dispositions), because said properties are not usually unambiguously bound to a specific set of spatiostructural properties [70]. According to [3], certain groups of social insects possess this characteristic uncoupling of disposition (in his example, reproductive disposition) from a specific set of spatiostructural properties (e.g., development of morphoanatomical components associated with wings). In cases where researchers want to find clearly specifiable spatiostructural recognition criteria for

functionally defined biological entities, they must conduct experiments that can clearly identify a particular functionally defined spatiostructural entity (e.g., through physiological experiments). This situation leads to Wilson's third issue.

One very important thing to highlight is that most of the time, researchers do not have access to these experiments, and, when they try to specify recognition criteria for functionally defined spatiostructural entities, they are restricted to differential diagnosis and comparative approaches coupled with conclusions based on analogy ([69]). As a compound, non-deductive, type of reasoning [71], [72], analogical reasoning involves inferential moves that spams different frames of reference – i.e., it involves processes of matching and mapping between different conceptual domains – which can lead to logical inconsistencies when misapplied. Given the complexity of the inferential form of analogy, and considering the lack of a clear logical structure for this type of reasoning, recognizing 'castes' by a set of functionally defined spatiostructural entities can potentially lead to ontological errors and creation of several unnested categories.

Conclusion

Using a well-defined framework to organize 'colonial' and 'caste' concepts enables coherent representation of several dimensions of biological reality. Although future efforts are needed to address specific issues in 'colony' and 'caste' modeling in the bio-ontologies, the use of a logically aligned framework is a major step forward in dealing with much of the inconsistency in demarcation of fundamental entities in the study of social insects.

Acknowledgements

The authors are deeply indebted to Alexandre Casadei Ferreira, Amanda M. Dias, Natalia Ladino, Mila F. O. Martins, Paloma L. Andrade, and Tainara T. Jory, who provided suggestions and discussions on initial versions of this manuscript. We are deeply grateful to Lars Vogt, who pointed out several inconsistencies in previous versions of this work. TSRS was funded by CAPES and Fundação Araucária (process number 88887.354448/2019-00), and by the Environment and Conservation Fund from the Government of the Hong Kong Special Administrative Region, under the 'Environmental Research, Technology Demonstration and Conference Projects' funding scheme – Project number ECF 137/2020.

539

References

[1] Costa JT, Fitzgerald TD. Developments in social terminology: semantic battles in a conceptual war. Trends Ecol Evol. 1996; 11(7): 285-289.

[2] Costa JT, Fitzgerald TD. Social terminology revisited: Where are we ten years later?. In: Annales Zoologici Fennici 42(6). Finnish Zoological and Botanical Publishing Board; 2005. pp. 559-564.

[3] Peeters C. Convergent evolution of wingless reproductives across all subfamilies of ants, and sporadic loss of winged queens (Hymenoptera: Formicidae). Myrmecol News. 2012; 16: 75-91.

[4] Neco LC, Japyassú HF, El-Hani CN, Châline N. From classificatory to quantitative concepts in the study of sociality in animals: an epistemological view. Biol Theory. 2018; 13(3): 180-189.

[5] Silva TSR, Feitosa RM. On titles and royalty: a terminological discussion over castes in myrmecology. Insectes Soc. 2019; 66(1): 25-35.

[6] Sumner S, Bell E, Taylor D. A molecular concept of caste in insect societies. Curr Opin Insect Sci. 2018; 25: 42-50.

[7] Crespi BJ, Yanega D. The definition of eusociality. Behav Ecol. 1995; 6(1): 109-115.

[8] Dew RM, Tierney SM, Schwarz MP. Social evolution and casteless societies: needs for new terminology and a new evolutionary focus. Insectes Soc. 2016; 63(1): 5-14.

[9] Epstein M. Nomenclature, terminology and language. Bionomina. 2012; 5(1): 1-56.

[10] Faber P. Frames as a framework for terminology. In: Kockaert HJ, Steurs F. Handbook of terminology 1(14). John Benjamins Publishing Company; 2015; pp. 14-33.

[11] Faber P, León-Araúz P. Bridging the Gap between Description and Standardization: a Frame-based Version of NATO Glossaries. In: Margalitadze T, Meladze G. Proceedings of the XVII EURALEX International Congress. Lexicography and

Linguistic Diversity. Tbilisi: Ivane Javakhishvili Tbilisi University Press; 2016. Pp. 499-508.

[12] Michener CD. The social behavior of the bees. Cambridge: Harvard University Press; 1974.

[13] Wilson EO. Sociobiology: The New Synthesis. Cambridge: Harvard University Press; 1975.

[14] Ryle G. The Concept of Mind. 60th Anniversary Edition. London: Routledge; 2009.

[15] Peeters C. The occurrence of sexual reproduction among ant workers. Biol J Linn Soc. 1991; 44(2): 141-152.

[16] Peeters C. Castas: homología y analogía en la forma y función. In: Fernández F, Guerrero RJ, Delsinne T. Hormigas de Colombia. Bogotá: Universidad Nacional Colombia; 2019. Pp. 159–164.

[17] Peeters C, Crozier RH. Caste and reproduction in ants: not all mated egg-layers are "queens". Psyche. 1988; 95(3-4): 283-288.

[18] Hölldobler B, Wilson EO. The ants. Cambridge: Harvard University Press; 1990.

[19] Buschinger A, Winter U. (1976). Funktionelle Monogynie bei der Gastameise *Formicoxenus nitidulus* (NYL.) (Hym., Form.). Insectes Soc. 1976; 23: 549-558.

[20] Buschinger A, Winter U. Echte Arbeiterinnen, fertile Arbeiterinnen und sterile Wirtsweibchen in Völkern der dulotischen Ameise *Harpagoxenus sublaevis* (NYL.) (Hym., Form.). Insectes Soc. 1978; 25: 63-78.

[21] Tanney J. Rethinking Ryle: A Critical Discussion of the Concept of Mind. In: Ryle G. The Concept of Mind. 60th Anniversary Edition. London: Routledge; 2009. Pp. ix-lvii.

[22] Magidor, O Category Mistakes. 2019 Jul 5 [cited 28 April 2021]. In: The Stanford Encyclopedia of Philosophy (Winter 2020 Edition) [Internet]. Available from: https://plato.stanford.edu/archives/win2020/entries/category-mistakes.

[23] Vogt L. Bona fideness of material entities and their boundaries. In: Davies R. Natural and artifactual objects in contemporary metaphysics: exercises in analytical ontology. London: Bloomsbury Academic; 2019. Pp. 103-120.

[24] Herbers JM. Watch your language! Racially loaded metaphors in scientific research. Bioscience. 2007; 57(2): 104-105.

[25] Herbers JM. Racist words in science. Bioscience. 2020; 70(11): 946-946.

[26] Sleigh C. Ant. Chicago: Chicago University Press, Reaktion Books; 2004.

[27] Cabré MT. Theories of terminology: Their description, prescription and explanation. Terminology. 2003; 9(2): 163-199.

[28] Vogt L. Levels and building blocks—toward a domain granularity framework for the life sciences. J Biomed Semant. 2019; 10(1): 1-29.

[29] Smith B. Basic Formal Ontology 2.0: Specification and user's guide. 2015 Jun 26 [cited 18 May 2021]. In: Github [Internet]. Available from: https://github.com/bfo-ontology/BFO/wiki.

[30] Smith B, Almeida M, Bona J, Brochhausen M, Ceusters W, et al. Basic formal ontology 2.0. 2015. [cited 18 May 2021]. In: Github [Internet]. Available: https://github.com/BFO-ontology/BFO2.

[31] Vogt L, Grobe P, Quast B, Bartolomaeus T. Fiat or bona fide boundary—a matter of granular perspective. PLoS One. 2012; 7(12): e48603.

[32] Strassmann JE, Queller DC. (2010). The social organism: congresses, parties, and committees. Evol Int J Org Evol. 2010; 64(3): 605-616.

[33] Bonabeau, E. Social insect colonies as complex adaptive systems. Ecosystems. 1998; 1(5): 437-443.

[34] Bonabeau E, Theraulaz G, Deneubourg JL, Franks NR, Rafelsberger O, Joly JL, Blanco S. A model for the emergence of pillars, walls and royal chambers in termite nests. Philos Trans Royal Soc Lond, B, Biol Sci. 1998; 353(1375): 1561-1576.

[35] Levin SA. 1998. Ecosystems and the biosphere as complex adaptive systems. Ecosystems. 1998; 1: 431–36.

[36] Mlot NJ, Tovey CA, Hu DL. Fire ants self-assemble into waterproof rafts to survive floods. Proc Natl Acad Sci. 2011; 108(19): 7669-7673.

[37] Phonekeo S, Mlot N, Monaenkova D, Hu DL., Tovey C. (2017). Fire ants perpetually rebuild sinking towers. R Soc Open Sci. 2017; 4(7): 170475.

[38] Franks NR. (1989). Thermoregulation in army ant bivouacs. Physiol Entomol. 1989; 14(4): 397-404.

[39] Ingarden R. (1983) Man and value. München: Philosophia Verlag GmbH. p. 184.

[40] Smith B, Brogaard B. (2003) Sixteen days. J Med Philos. 2003; 28: 45–78.

[41] Schmickl T, Karsai I. Integral feedback control is at the core of task allocation and resilience of insect societies. Proceedings of the National Academy of Sciences. 2018; 115(52): 13180-13185.

[42] Jones JC, Myerscough MR, Graham S, Oldroyd BP. Honey bee nest thermoregulation: diversity promotes stability. Science. 2004; 305(5682): 402-404.

[43] Smith B, Varzi AC. The niche. Noûs. 1999; 33(2): 214-238.

[44] Smith B, Varzi AC. Environmental metaphysics. In: Meixner U, Simons PM. Metaphysics in the Post-Metaphysical Age. Proceedings of the 22nd International Wittgenstein-Symposium. Vienna: Hölder-Pichler-Tempsky; 2001. Pp. 231-242.

[45] Casati R, Varzi AC. Holes and Other Superficialities. Cambridge: MIT Press; 1994.

[46] Bartholomew GA, Lighton JR, Feener Jr DH. Energetics of trail running, load carriage, and emigration in the column-raiding army ant *Eciton hamatum*. Physiol Zool. 1988; 61(1): 57-68.

[47] Oi CA, van Zweden JS, Oliveira RC, Van Oystaeyen A, Nascimento FS, Wenseleers T. The origin and evolution of social insect queen pheromones: novel hypotheses and outstanding problems. BioEssays. 2015;37(7): 808-821.

[48] Pacala SW, Gordon DM, Godfray HC. Effects of social group size on information transfer and task allocation. Evol Ecol. 1996; 10(2): 127-65.

[49] Modlmeier AP, Colman E, Hanks EM, Bringenberg R, Bansal S, Hughes DP. Ant colonies maintain social homeostasis in the face of decreased density. Elife. 2019; 8: e38473.

[50] Baracchi D, Cini A. A socio□spatial combined approach confirms a highly compartmentalised structure in honeybees. Ethology. 2014; 120(12): 1167-1176.

[51] Heyman Y, Shental N, Brandis A, Hefetz A, Feinerman O. Ants regulate colony spatial organization using multiple chemical road-signs. Nat Commun. 2017; 8(1): 1-1.

[52] Jandt JM, Dornhaus A. Spatial organization and division of labour in the bumblebee *Bombus impatiens*. Anim Behav. 2009; 77(3): 641-651.

[53] Mersch DP, Crespi A, Keller L. Tracking individuals shows spatial fidelity is a key regulator of ant social organization. Science. 2013; 340(6136):1090-1093.

[54] Powell S, Tschinkel WR. Ritualized conflict in *Odontomachus brunneus* and the generation of interaction-based task allocation: a new organizational mechanism in ants. Anim Behav. 1999; 58(5): 965-72.

[55] Robson SK, Bean K, Hansen J, Norling K, Rowe RJ, White D. Social and spatial organisation in colonies of a primitively eusocial wasp, *Ropalidia revolutionalis* (de Saussure) (Hymenoptera: Vespidae). Aust J Entomol. 2000; 39(1): 20-24.

[56] Seeley TD. Adaptive significance of the age polyethism schedule in honeybee colonies. Behav Ecol Sociobiol. 1982; 11(4): 287-293.

[57] Sendova-Franks AB, Franks NR. Spatial relationships within nests of the ant *Leptothorax unifasciatus* (Latr.) and their implications for the division of labour. Anim Behav. 1995; 50(1): 121-136.

[58] Londe S, Monnin T, Cornette R, Debat V, Fisher BL, Molet M. Phenotypic plasticity and modularity allow for the production of novel mosaic phenotypes in ants. Evodevo. 2015; 6(1): 1-5.

[59] Hartfelder K, Makert GR, Judice CC, Pereira GA, Santana WC, Dallacqua R, Bitondi MM. Physiological and genetic mechanisms underlying caste development, reproduction and division of labor in stingless bees. Apidologie. 2006; 37(2): 144-163.

[60] Peeters C, Molet M. Colonial reproduction and life histories. In: Lach L, Parr C, Abbott K. Ant Ecology. Oxford University Press; 2010. Pp. 159-176

[61] Trible W, Kronauer DJ. Hourglass model for developmental evolution of ant castes. Trends Ecol. Evol. 2021; 36(2): 100-103.

[62] Kuhn A, Darras H, Aron S. Phenotypic plasticity in an ant with strong caste–genotype association. Biol Lett. 2018; 14(1): 20170705.

[63] Vogt L, Grobe P, Quast B, Bartolomaeus T. Top-level categories of constitutively organized material entities-suggestions for a formal top-level ontology. PLoS One. 2011; 6(4): e18794.

[64] Craver CF, Bechtel W. Top-down causation without top-down causes. Biol Philos. 2007; 22: 547–63.

[65] Crozier RH, Pamilo P. Evolution of Social Insect ‘colonies’: Sex Allocation and Kin Selection. Oxford: Oxford University Press; 1996.

[66] Elsner D, Meusemann K, Korb J. Longevity and transposon defense, the case of termite reproductives. Proc Natl Acad Sci. 2018; 115(21): 5504-5509.

[67] Buttigieg PL, Pafilis E, Lewis SE, Schildhauer MP, Walls RL, Mungall CJ. The environment ontology in 2016: bridging domains with increased scope, semantic density, and interoperation. J. Biomed. Semantics. 2016; 7(1): 1-12.

[68] Walls R, Buttigieg PL. Semantically Defining Populations for 'Omics Research. Biodiversity Information Science and Standards. 2017; 1: e20435.

[69] Wilson EO. The evolution of caste systems in social insects. Proc Am Philos Soc. 1979; 123(4): 204-210.

[70] Vogt L, Mikó I, Bartolomaeus T. Anatomy and the type concept in biology show that ontologies must be adapted to the diagnostic needs of research. J. Biomed. Semantics. 2022; 13(1): 1-27.

[71] Minnameier G. Abduction, induction, and analogy – On the compound character of analogical inferences. In: Carnielli W, Magnani L, Pizzi C. Model-Based Reasoning in Science and Technology: Abduction, Logic, and Computational Discovery. Heidelberg: Springer Verlag GmbH; 2010. Pp. 107–119.

[72] Peirce CS. 1896. Lessons from the history of science. In: Hartshorne C, Weiss P, Burks A. Peirce CS. Collected Papers of Charles Sanders Peirce. Cambridge: Harvard University Press; 1958.

Figures

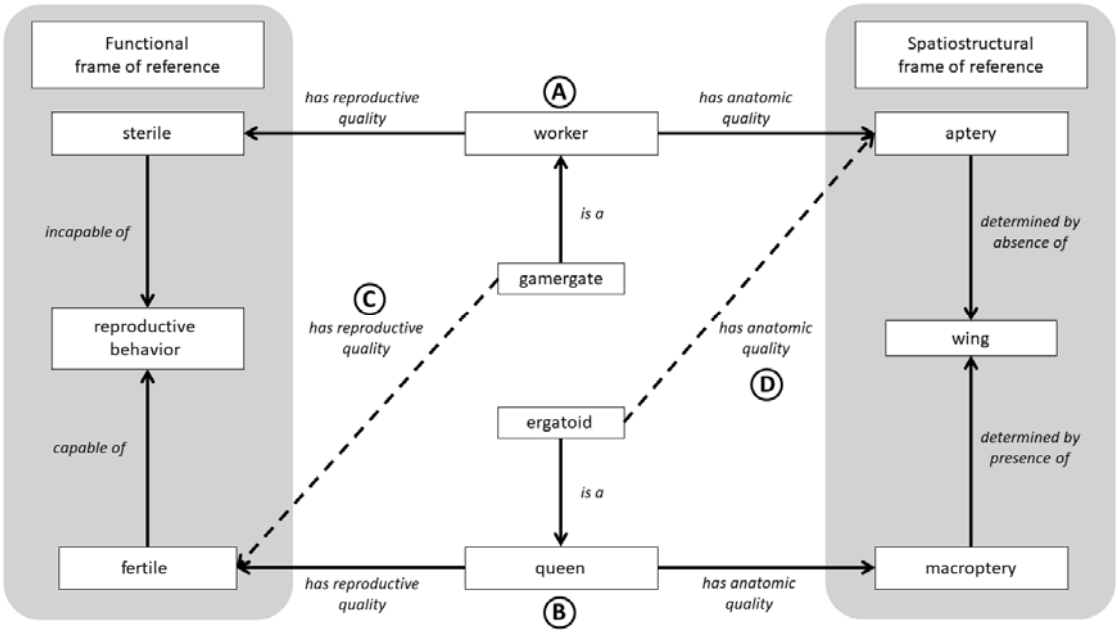

Figure 1. Example of category mistake in two 'caste' concepts applied in myrmecology: gamergate and ergatoid. Equivocal relations leading to category mistakes are marked as dotted lines. In this example, most relational properties are not necessarily compliant with formal ontologies (e.g., *incapable of*, *determined by presence of*, *has anatomic quality*) and are used only for didactic purposes.

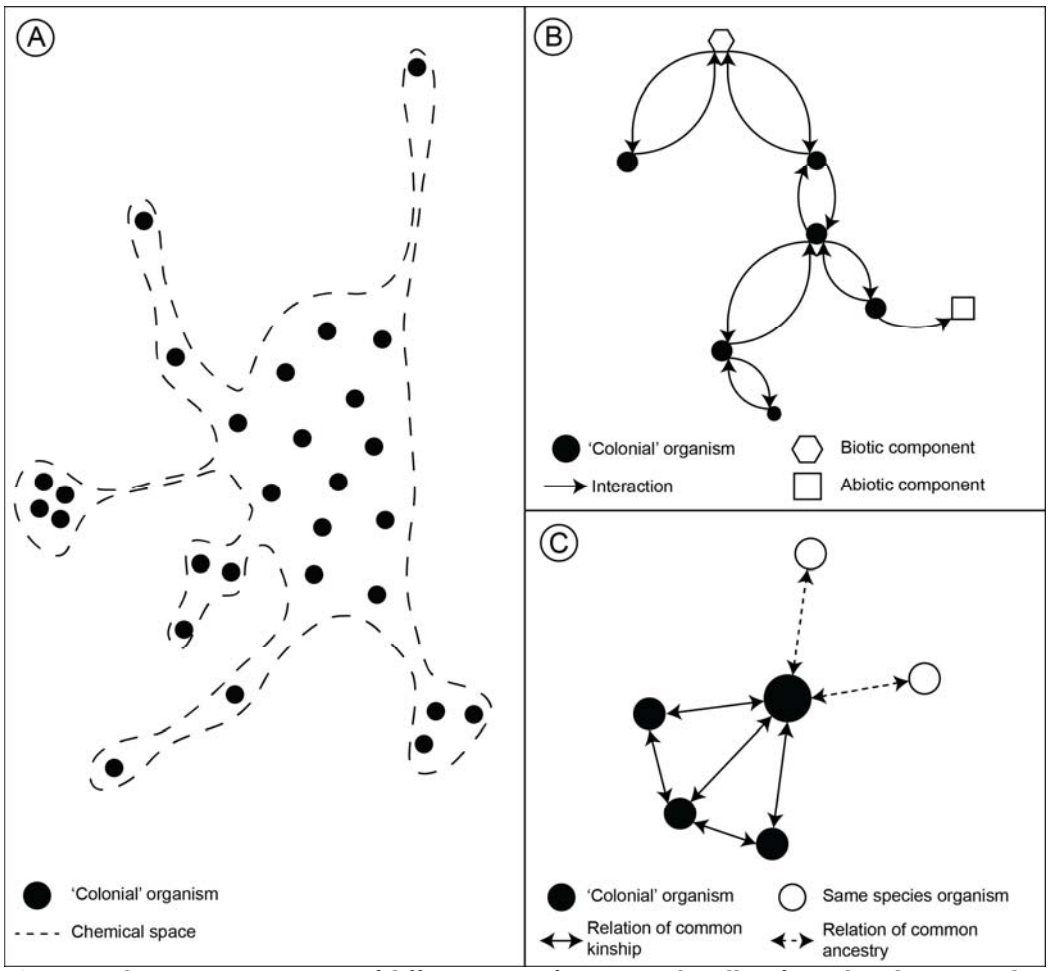

Figure 2 Schematic representation of different types of properties that allow for 'colonial' connectedness and their respective frames of reference. A. 'Colonial' connectedness through causal unity via internal physical forces (i.e., chemical space); B. 'Colonial' connectedness through causal unity via physiological dispositions (i.e., ecological interactions); C. 'Colonial' connectedness through causal unity via historical origin of its parts (i.e., relations of common kinship).