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

Thiago S.R. Silva[1,2]

[1]Departamento de Zoologia, Universidade Federal do Paraná, Curitiba, Brazil.
[2]School of Biological Sciences, The University of Hong Kong, Hong Kong SAR, China.

Contributing authors: 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. I 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, I briefly discuss some limitations in current approaches to demarcate subcategories of 'castes'.

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

1 Introduction

Terminological standardization has been a matter of debate among research groups that study insect societies for several years (Costa & Fitzgerald 1996, 2005; Peeters 2012; Neco et al. 2018; Silva & Feitosa 2019a; Sumner et al. 2018). According to several authors (e.g. Costa & Fitzgerald 2005; Crespi & Yanega 1995; Dew et al. 2016; Sumner et al. 2018), 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 more straightforward communication, whether textually or orally, and (iii) to provide a framework for unified concept representation and organization, which precedes implementations of formal specifications aimed at (meta)data interoperability, ultimately enhancing (meta)data comparability and reproducibility (Wilkinson et al. 2016). Aiming for concept clarity is an important first step for the 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 (Epstein 2012; Faber 2015; Faber & León-Araúz 2016). While the abovementioned 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 (Michener 1974; Wilson 1975; Dew et al. 2016; Silva & Feitosa 2019a).

Although employing different frames of reference during concept representation can be considered the best approach for demarcating biological entities, researchers can equivocally define concepts through relations that violate one or more classification criteria in a given conceptual scheme, leading to inconsistent classifications. These inconsistencies can happen when enunciators try to classify a set of entities using a *more-than-one* set of independent and not disjoint properties. Examples of inconsistent classifications in 'caste' representation include gamergates and ergatoids, as commonly understood in myrmecology (Peeters 1991; Peeters & Crozier 1988; Peeters 2019) (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) (Hölldobler & Wilson 1990; Buschinger & Winter 1976; 1978; Peeters & Crozier 1988; Peeters 2019). 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).

This work aims 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. I will argue that both concepts can be demarcated under two common frames of reference used in biological sciences, according to Vogt (2019a): functional and historical/evolutionary. will adjust and refine previous propositions for a 'caste' ontology while addressing some misconceptions made by Silva & Feitosa (2019a). Initially, I will argue that some fundamental properties are commonly evoked by researchers when they try to identify individual 'colonies' of social insects, which, in turn, are used to partition them as natural units. After that, I will argue that all individual 'colonies' share defining properties among them, establishing 'colony' as a natural kind, with each individual 'colony' (e.g. a *Coptotermes* 'colony') instantiating the class 'colony'. I will then argue that sets of

organisms have shared properties that allow researchers to establish 'castes' as natural kinds, with each individual 'caste' (e.g. a *Vespula* queen) instantiating a certain 'caste' class (e.g. the queen 'caste'). Lastly, I will provide a brief account of the limitations of current approaches to demarcate subcategories of 'castes'.

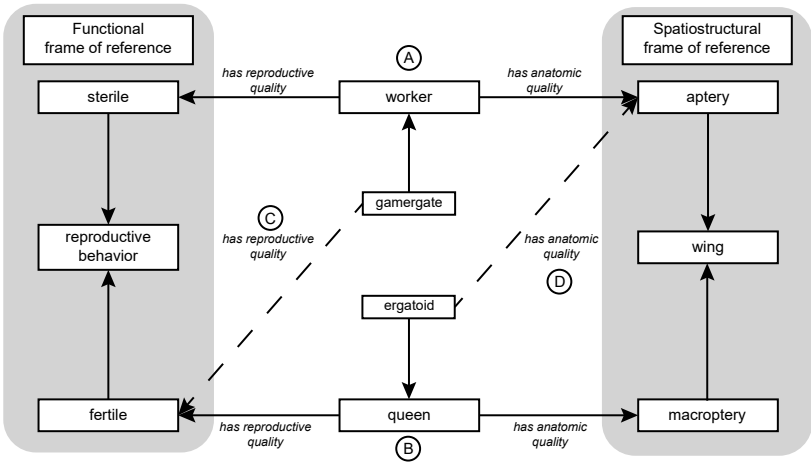

Fig. 1 Example of 'inconsistent classification' in two 'caste' concepts applied in myrmecology: gamergate and ergatoid. Violations of classification criteria are marked as dotted lines. In this example, most relational properties are not necessarily compliant with formal ontologies (e.g., *incapable of*, *determined by presence of*, *has anatomic quality*) and are used only for didactic purposes.

2 Methods

2.1 Specialized terms

In this work, I 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 Herbers (2007; 2020) and Sleigh (2004)]. I understand that reevaluating the linguistic component of terminological units – i.e., multidimensional units composed of linguistic, cognitive, and situational components (Cabré 2003) – is essential for the development of a more inclusive and diverse science. I 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, I 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 I change an oppressive-laden name by another.

2.2 Frames of reference and relations of causal unity

I 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, I 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 (Vogt 2019b). To deal with these issues, I first defined 'colonies' as natural units via specific types of causal unities, which, in turn, are respectively connected to specific frames of reference (Smith et al. 2015; Vogt 2019a). Then, I demarcated both 'colonies' and 'castes' through certain properties shared by their respective instances.

To specifically address the mind-independence of non-structural boundaries of 'colonies', I followed the examples of functional and historical/evolutionary units provided by Vogt et al. (2012), which can be divided into (i) dispositions independent of morphogenesis and (ii) dispositions involving morphogenesis for the functional units, 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 historical/evolutionary units.

3 Results

3.1 'Colonies' of social insects as unified causal systems

Insect 'colonies' are considered 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 (Strassmann & Queller 2010). 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 (Bonabeau 1998; Bonabeau et al. 1998) – i.e., systems where 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 (Levin 1998). 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., Bonabeau et al. (1998), Mlot et al. (2011), Phonekeo et al. (2017), and Franks (1989)], while emergence explanations of historical/evolutionary properties are non-compliant with the model.

To address these limitations, I work under the premise that 'colonies' are causally relatively isolated entities [i.e., they are unified causal systems isolated from their surroundings; Ingarden (1983) and Smith & Brogaard (2003)] 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, five types of causal unities 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 (Smith et al. 2015; Vogt 2019a). 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) (Vogt 2019a).

Ideally, 'colonies' of social insects are maintained as causally connected units thanks to some properties that allow them to achieve some degree of internal homeostasis (Schmickl & Karsai 2018). The main properties that enable 'colonial' connectedness are chemical space, self-organization, and resilience. These properties can be linked to specific frames of reference and the parts that enact them can be further partitioned in their respective perspectives. In the following sections, I will show how chemical space, self-organization, and resilience can be connected to functional and historical/evolutionary frames of reference, respectively.

3.1.1 Functional frame of reference and functional 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 (Jones et al. 2004). 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 Smith et al. (2015), since they are not comprised solely of an enclosing physical structure, but rather are considered as niche tokens (Smith & Varzi 1999) in the form of double hole structures [(Smith & Varzi 2001; for holes see Casati & Varzi (1994)]. 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, the bivouac temperature at its core is maintained stable and higher than the environment temperature (Franks 1989), through higher respiratory rates of workers when they are gathered (Bartholomew et al 1988); this property allows for fast development of the immature stages and 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 chemical connections 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 functionally 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 functional (as well as historical/evolutionary; see discussion in Results 3.1.2) connection of two or more organisms of the aggregate. The process of pheromone production and recognition is considered to provide functional connection because, although each component involved in the process is structurally bound (normally in the form of weak interactions) to the immediate next component in the chain of events that corresponds to the production-recognition process, ultimately the result of this process is functional, i.e. the output of the chain is a set of behaviors of acceptance/aggression towards members of the 'colony'.

There are some studies discussing the importance of chemical distance and its relation with the 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 (Oi et al. 2015). Chemical signatures are a major force promoting functional connectedness of insect aggregations; 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'.

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 nest), some members of a 'colony' are spatially separated from other members by disconnected nests. Nonetheless, they share the same chemical space and, thus, are functionally connected.

It is important to consider, however, that insect 'colonies' are dynamic systems that vary in their functional connectedness through time. This is especially true when a single individual (or a group of individuals that are a subset of the 'colony') loses their connectedness with the whole during a given time – i.e. when the individual leaves the pheromone space of the colony. This limitation will be accounted for when I discuss resilience and how it promotes 'colony' connectedness (see Results 3.2.1).

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 sub-optimal 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 (Pacala et al. 1996).

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 in the vicinity 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 feedback: new arrivals at the food source are considered as 'births', while the abandonment of food sources is considered as 'deaths'. The recruitment ceases when the system reaches a 'stable population size', meaning that the 'birth rate' equals 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 Modlmeier et al. (2019), 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 (Baracchi & Cini 2014; Heyman et al 2017; Jandt & Dornhaus 2009; Mersch et al 2013; Powell & Tschinkel 1999; Robson et al. 2000; Seeley 1982). These movement zones [spatial fidelity zones according to Sendova-Franks & Franks (1995)] 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 constrain 'colonial' connectedness (Modlmeier et al. 2019). Hence, changes in the 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.

3.1.2 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 (Strassman & Queller 2010). 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 grow 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).

3.2 'Colonies' and 'castes' of social insects as natural kinds

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 a response to intrinsic and/or extrinsic stimuli. From that, we can assume that 'castes' are the 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's' 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, each organism that differentiates into a certain 'caste' will contribute to the 'colonial' stability through certain specific dispositions. However, contrary to 'colonies', 'castes' cannot be considered natural units themselves, since they are already considered unified causal systems given their organism-specific properties.

Despite that, different sets of organisms living in a 'colony' share some properties that allow researchers to demarcate specific classes of organisms living in a 'colony'. These properties can contribute or not to the overall stability of a 'colony'; nonetheless,

they are specifically bound to certain sets of organisms. These shared properties can also be found among sets of organisms from different colonies. Given this specificity in shared properties among sets of organisms from different 'colonies', researchers often assume that these sets of organisms form natural groups in the form of 'castes'.

Hence, researchers tend to agree that 'colony' and 'caste' are terms that carry meaning in the form of direct representations of entities that exist in a given biological reality. In this case, the entities that exist in reality are natural kinds, while the cognitive representations of those entities are the concepts, with the names used to evoke these representations the kind terms (Vogt et al. 2022). The concepts of 'colonies' and 'castes' are commonly established as classes and defined intensionally through a *genus differentia* definition (e.g. a queen is a female that has a reproductive disposition). However, in these particular cases, intensional definitions are not sufficient, since they only provide ontological knowledge about the entities in question - i.e. they offer transparency regarding the meaning of its associated kind term -, while diagnostic knowledge is not covered - i.e. researchers normally have difficulty in correctly applying a term to a given real entity. Adding empirical recognition criteria to 'colonies' and 'castes' will further help researchers unambiguously reference any of those kind terms.

Table 1 Aristotelian definition, in the form of *genus differentia*, of 'colony' and 'caste' entities.

Entity	*Genus differentia* definition
'Colony'	A group of organisms that have the disposition for self-organization and have the historical origins of its parts through relations of common kinship.
'Caste'	An individual who has the disposition to perform a specific function as a realization of a specific developmental fate.

3.2.1 Definition and recognition criteria of 'colony'

The Aristotelian definition of 'colony' would comprise both the disposition of a group of individuals to self-organize and the historical origin of its parts through relations of common kinship (Table 1). However, its definition is not sufficient for a researcher to identify the entity, with more information being required on how a 'colony' looks like (Table 2).

Since 'colonies' are not defined by any spatio-structural property, just by functional and historical/evolutionary properties of its parts, it is not possible to specify any structural recognition criterion. This poses a problem in recognizing a member of a certain 'colony' since its functional property - i.e. disposition for self-organization - can only be observed under certain behavioral essays, which cannot be elaborated every single time a researcher needs to identify a 'colony'. Additionally, historical properties of 'colonies' - i.e. relations of common kinship - cannot be directly observed. Additionally, not all instances of the class 'colony' will have exactly all the defining properties of it. In cases like this, researchers cannot apply essentialistic class concepts for the recognition of instances of the class and must resort to some vague notion of type (Vogt et al. 2022). In doing so, they need to meet certain demands for recognizing

Table 2 Diagnostic recognition criteria of 'colony'

Method	Recognition criteria
Gas chromatography coupled with mass spectrometric detection	*Members of 'colony' synthesize a 'colony'-specific multi-component aggregation pheromone*
Gas chromatography coupled with mass spectrometric detection	*Members of 'colony' encode colony-specific information in the proportions of individual hydrocarbons within a single homologous series*
Gas chromatography coupled with mass spectrometric detection	*Members of 'colony' synthesize a combination of specific compounds principally in one-carbon chain length or in limited homologous series*
Attraction/avoidance behavior analysis	*Members of 'colony' attract or avoid other individuals from the same species*

class membership of the instances that belong to this vague notion of 'colony' - which, instead of being treated as an essentialistic class, is now considered a diagnostic concept in the form of a *cluster class*.

Hence, the instances that need to be categorized must reach a *minimum quorum* of properties to be ascribed to the class 'colony'. In the case of 'colonies', a minimum quorum of three properties is considered to be sufficient (Table 3).

Table 3 Seven particulars (A-G) with the properties listed above used to recognize the entity 'colony'

Instance	A	B	C	D	E	F	G
Property 1	X	X	X		X	X	X
Property 2				X	X		X
Property 3	X	X	X			X	
Property 4	X	X			X		X
cluster class membership (*minimum quorum* 3)	yes	yes	no	no	yes	no	yes

3.2.2 Definition and recognition criteria of 'castes'

The Aristotelian definition of 'caste' relates to the disposition of an individual to perform a specific function as a realization of a specific developmental fate (Table 1). Similarly to 'colony', this definition is not sufficient for the identification of the entity and more information is required so researchers can answer the question of how a 'caste' looks like (Table 4).

'Castes' are determined by several developmental changes induced by both genetic and environmental constraints, transforming their spatiostructurally makeup at the end of the process. Since 'caste' determination depends on several genetically determined parts, and those parts are intrinsically dependent on 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 is causally linked and coordinated with one another. 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 (Londe et al 2015), physiological determinants (Hartfelder et al. 2006), and/or environmental constraints (Peeters & Molet 2010). Hence, each component on a specific level of organization has differential participation in the determination of particular 'castes'. Those components can (a) individually and directly influence the determination of the 'caste', or (b) synergistically 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'.

Table 4 Diagnostic recognition criteria of 'caste'

Method	Recognition criteria
Reflected light, stereomicroscope	*Individual has differential physical characteristics, such as size, body proportions, presence of specialized structures, color patterns, setae and relief density, and head and mandible size and shape, when compared to other individuals from the same 'colony'*
Bright field, microscope	*Individual has differential stages of development of reproductive organs when compared to other individuals from the same 'colony'*
Behavioral experiment, pheromone manipulation	*Individual has differential distribution on space of 'colony', with certain individuals being concentrated in specific areas occupied by the 'colony'*
RNAseq and microarray analysis	*Individual has differential gene expression, representing (or not) the ancestral gene regulatory network*
Whole-Genome Bisulfite Sequencing	*Individual has differential DNA methylation patterns in certain regions compared to other individuals from the* same 'colony'

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 (Trible & Kronauer 2020). 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 (Kuhn et al. 2018). Hence, the determination of 'castes' can be directed at the genotype level in some cases, while being directed at the metabolic or environmental level in other cases. 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.

'Castes' are composed of particular parts that have proven in the past that they have developed in concert, 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. Hence, similar to 'colonies', the instances that need to be categorized must reach a minimum quorum of properties

to be ascribed to the class 'caste'. In the case of 'caste', a minimum quorum of three properties is considered to be sufficient (Table 5).

Table 5 Seven particulars (A-G) with the properties listed above used to recognize the entity 'caste'

Instance	A	B	C	D	E	F	G
Property 1	X	X	X			X	X
Property 2				X	X		X
Property 3	X	X				X	
Property 4	X	X			X		X
Property 5		X	X		X		X
cluster class membership (*minimum quorum* 3)	yes	yes	no	no	yes	no	yes

4 Discussion

4.1 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 data communication and management (Vogt et al. 2011). However, since several ontologies have been developed with a particular practical purpose in mind (Silva & Feitosa 2019a), with a focus on the definition of 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 semantic 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 representations 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 establish 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 (Craver & Bechtel 2007) 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 (Crozier & Pamilo 1996; Elsner et al. 2018).

In terms of ontology design, recent efforts bridged important gaps in suggesting additional ontology relations (Buttigieg et al. 2016; Walls & Buttigieg 2017). 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 its respective frame. This is somewhat problematic, especially if we are concerned with comparability and compatibility of different ontologies. As highlighted by Vogt et al. (2012), 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. One could also add, through class axioms, additional properties for each class, which in turn can be used to identify additional implicit class hierarchies by a reasoner.

4.2 Current limitations for 'caste' subcategories

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' (Silva & Feitosa 2019a). 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 classify '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 the meaning of the term. However, insofar, attempts to split 'castes' into subcategories have not been entirely successful. One prominent example refers to the concepts of 'physical castes' and 'temporal castes' proposed by Wilson (1979).

According to Wilson's proposition, a 'caste' would be an ensemble of colony members that specialize in particular tasks for prolonged periods, being typically (but not necessarily) distinguished by other genetic, anatomical, or physiological traits. He further categorizes 'castes' into 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, the translation of Wilson's proposition to bio-ontologies has some limitations: (i) ambiguous defining criteria of 'castes', which in turn are used to derive the defining criteria of the two proposed subcategories; and (ii) creation of several unnested categories.

Wilson's first issue directly stems from the lack of consideration of the distinct frames of reference used to define the recognition criteria for the class being demarcated. Wilson considers 'castes' as object aggregates in the sense of an organism as an aggregate of cells and molecules 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 (Vogt et al. 2022). According to Peeters (2012), 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 specifiable spatiostructural recognition criteria for functionally defined biological entities, they must conduct experiments that can identify a particular functionally defined spatiostructural entity (e.g., through physiological experiments). Alternatively, researchers could create a context in which a reaction is triggered that realizes the disposition and the reaction has directly or indirectly observable consequences, which can be used as recognition criteria. This leads to Wilson's "second 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 (Vogt et al. 2022). As a compound, non-deductive, type of reasoning (Minnameier 2010; Peirce 1896: 28p, CP 1.65), analogical reasoning involves inferential moves that spam 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. However, despite the complexity of the inferential form of analogy, one can recognize 'castes' by a set of functionally defined spatiostructural entities without potentially leading to ontological errors. Additionally, in terms of ontology design, unnestedness does not necessarily lead to errors; rather, the strict adherence to single-inheritance hierarchies when dealing with compositional classes can potentially lead to errors. In this sense, Wilson's "third issue" serves more as a cautionary note in design choice rather than an issue in ontological requirements. As a solution for the single-inheritance hierarchy question, researchers can account for additional class-subclass relationships by specifying within class axioms some properties that its instances necessarily have to possess and then use these axioms to infer subclasses. This way, it is possible to have a single-inheritance ontology - in which each class has only one direct parent class - and derive all implicit class-subclass relationships through reasoning.

5 Conclusion

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

6 Acknowledgments

The author is 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. I am deeply grateful to Lars Vogt, who pointed out several inconsistencies and misconceptions in previous versions of this work and proposed solutions to conflicts raised in the discussion.

Authors' contributions Conceptualization: TSRS; Methodology: TSRS; Visualization: TSRS; Writing – original draft: TSRS; Writing – review & editing: TSRS.

Funding 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.

Data Availability All data is available in the manuscript.

7 Declarations

Competing interests The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals, performed by any of the authors.

Consent to participate All authors agree to participate in this research project.

Consent for Publication All authors approve the publication of the manuscript.

8 References

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

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

Bonabeau, E (1998) Social insect colonies as complex adaptive systems. Ecosystems 1(5): 437–443.

Bonabeau E, Theraulaz G, Deneubourg JL, Franks NR, Rafelsberger O, Joly JL, Blanco S (1998) A model for the emergence of pillars, walls and royal chambers in termite nests. Philos T Roy Soc B 353(1375): 1561–1576.

Buschinger A, Winter U (1976) Funktionelle Monogynie bei der Gastameise *Formicoxenus nitidulus* (NYL.) (Hym., Form.). Insect Soc 23: 549–558.

Buschinger A, Winter U (1978) Echte Arbeiterinnen, fertile Arbeiterinnen und sterile Wirtsweibchen in Völkern der dulotischen Ameise *Harpagoxenus sublaevis* (NYL.) (Hym., Form.). Insect Soc 25: 63–78.

Buttigieg PL, Pafilis E, Lewis SE, Schildhauer MP, Walls RL, Mungall CJ (2016) The environment ontology in 2016: bridging domains with increased scope, semantic density, and interoperation. J Biomed Semant 7(1): 57.

Cabré MT (2003) Theories of terminology: Their description, prescription and explanation. Terminology 9(2): 163–199.

Casati R, Varzi AC (1994) Holes and Other Superficialities. MIT Press, Cambridge, MA.

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

Costa JT, Fitzgerald TD (2005) Social terminology revisited: Where are we ten years later?. Ann Zool Fenn 42(6): 559–564.

Craver CF, Bechtel W (2007) Top-down causation without top-down causes. Biol Philos 22:547–63.

Crespi BJ, Yanega D (1995) The definition of eusociality. Behav Ecol 6(1): 109–115.

Crozier RH, Pamilo P (1996) Evolution of Social Insect Colonies: Sex Allocation and Kin Selection. Oxford University Press, Oxford.

Dew RM, Tierney SM, Schwarz MP (2016) Social evolution and casteless societies: needs for new terminology and a new evolutionary focus. Insect Soc 63(1): 5–14.

Elsner D, Meusemann K, Korb J (2018) Longevity and transposon defense, the case of termite reproductives. P Natl Acad Sci USA 115: 5504–5509.

Epstein M (2012) Nomenclature, terminology and language. Bionomina 5(1): 1–56.

Faber P (2015) Frames as a framework for terminology. In: Kockaert HJ, Steurs F (eds) Handbook of terminology, 1st vol. John Benjamins, Amsterdam, pp 14–33.

Faber P, León-Araúz P (2016) Bridging the Gap between Description and Standardization: a Frame-based Version of NATO Glossaries. In: Margalitadze T, Meladze G (eds) Proceedings of the XVII EURALEX International Congress: Lexicography and Linguistic Diversity. Ivane Javakhishvili Tbilisi University Press, Tbilisi, pp 499–508.

Franks NR (1989) Thermoregulation in army ant bivouacs. Physiol Entomol 14(4): 397–404.

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

Herbers JM (2007) Watch your language! Racially loaded metaphors in scientific research. BioScience 57(2): 104–105.

Herbers JM (2020) Racist words in science. BioScience 70(11): 946–946.

Heyman Y, Shental N, Brandis A, Hefetz A, Feinerman O (2017) Ants regulate colony spatial organization using multiple chemical road-signs. Nat Commun 8(1): 1–11.

Hölldobler B, Wilson EO (1990) The ants. Harvard University Press, Cambridge.

Ingarden R (1983) Man and value. Philosophia Verlag GmbH, München.

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

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

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

Levin SA (1998) Ecosystems and the biosphere as complex adaptive systems. Ecosystems 1: 431–436.

Londe S, Monnin T, Cornette R, Debat V, Fisher BL, Molet M (2015). Phenotypic plasticity and modularity allow for the production of novel mosaic phenotypes in ants. Evodevo 6(1): 1–15.

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

Michener CD (1974) The social behavior of the bees. Harvard University Press, Cambridge.

Minnameier G (2010) Abduction, induction, and analogy – On the compound character of analogical inferences. In: Carnielli W, Magnani L, Pizzi C (eds) Model-Based Reasoning in Science and Technology: Abduction, Logic, and Computational Discovery. Springer, Heidelberg, pp. 107–119.

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

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

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

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

Pacala SW, Gordon DM, Godfray HCJ (1996) Effects of social group size on information transfer and task allocation. Evol Ecol 10: 127–165.

Peeters C (1991) The occurrence of sexual reproduction among ant workers. Biol J Linn Soc 44(2): 141–152.

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

Peeters C (2019) Castas: homología y analogía en la forma y función. In: Fernández F, Guerrrero RJ, Delsinne T. Hormigas de Colombia. Universidad Nacional Colombia, Bogota, pp. 159–164.

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

Peeters C, Molet M (2010) Colonial reproduction and life histories. In: Lach L, Parr C, Abbott K. Ant Ecology. Oxford University Press, Oxford, pp 159–176.

Peirce CS (1896) Lessons from the history of science. In: Hartshorne C, Weiss P (1–6 vols., 1931–1935), Burks A (7–8 vols., 1958) (eds.) Collected Papers of Charles Sanders Peirce. Harvard University Press, Cambridge, pp. 43–125.

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

Powell S, Tschinkel WR (1999) Ritualized conflict in *Odontomachus brunneus* and the generation of interaction-based task allocation: a new organizational mechanism in ants. Anim Behav 58(5): 965–972.

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

Schmickl T, Karsai I (2018) Integral feedback control is at the core of task allocation and resilience of insect societies. P Natl Acad Sci 115(52): 13180–13185.

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

Sendova-Franks AB, Franks NR (1995) Spatial relationships within nests of the antLeptothorax unifasciatus (Latr.) and their implications for the division of labour. Anim Behav 50(1): 121–136.

Silva TSR, Feitosa RM (2019) On titles and royalty: a terminological discussion over castes in myrmecology. Insect Soc 66(1): 25–35.

Sleigh C (2004) Ant. Chicago University Press, Chicago.

Smith B, Almeida M, Bona J, Brochhausen M, Ceusters W, et al (2015) Basic Formal Ontology 2.0: Specification and user's guide. https://github.com/bfo-ontology/BFO/wiki. Accessed 22 May 2023.

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

Smith B, Varzi AC (1999) The niche. Noûs 33(2): 214–238.

Smith B, Varzi AC (2001) Environmental metaphysics. In: Meixner U, Simons PM (eds.) Metaphysics in the Post-Metaphysical Age: Proceedings of the 22nd International Wittgenstein-Symposium. Hölder-Pichler-Tempsky, Vienna, pp 231–242.

Sumner S, Bell E, Taylor D (2018) A molecular concept of caste in insect societies. Curr Opin Insect Sci 25: 42–50.

Strassmann JE, Queller DC (2007) Insect societies as divided organisms: the complexities of purpose and cross-purpose. P Natl Acad Sci 104(1): 8619–8626.

Trible W, Kronauer DJ (2020) Hourglass Model for Developmental Evolution of Ant Castes. Trends Ecol Evol 36(2): 100–103.

Vogt L (2019a) Bona fideness of material entities and their boundaries. In: Davies R (ed.) Natural and artifactual objects in contemporary metaphysics: exercises in analytical ontology. Bloomsbury Academic, London, pp 103–120.

Vogt L (2019b) Levels and building blocks—toward a domain granularity framework for the life sciences. J Biomed Semant 10(1): 4.

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

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

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

Walls R, Buttigieg PL (2017) Semantically Defining Populations for 'Omics Research. Biodivers Inf Sci Stand 1: e20435.

Wilkinson M, Dumontier M, Aalbersberg IJ, et al (2016) The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3: 160018. https://doi.org/10.1038/sdata.2016.18

Wilson EO (1975) Sociobiology: The New Synthesis. Harvard University Press, Cambridge.

Wilson EO (1979) The evolution of caste systems in social insects. P Am Philos Soc 123(4): 204–210.