

# THE **Festivus**

Vol. 55(2)

May 2023



**A new deep water *Fasciolaria*  
*Sciteconus* from Southeast Africa  
Amazing Andaman Sea *Murex*  
Iconography of Brazilian Conidae**

Quarterly Publication of the San Diego Shell Club

## The Completeness of Thematic Collections - a Venn Diagram-Based Conceptual Approach

Cesare Brizio

World Biodiversity Association, Museo Civico di Storia  
Naturale di Verona, Lungadige Porta Vittoria, 9 - 37129 Verona, Italy  
[briziocesare@gmail.com](mailto:briziocesare@gmail.com)

**ABSTRACT** Multiple factors influence the degree of completeness of any thematic collection, such as a biological collection encompassing any natural group, but even any collection of man-made items (including art collections and object collections) that encompass any specific theme. The concept of “completeness” itself is both subjective, inasmuch as it may refer to any purpose unilaterally defined by the collector, and objective, inasmuch as that purpose includes an exhaustive coverage of the actual diversity inherent in the chosen group. In a general sense, also the factors (or “conditions”) that exert their influence on the collection process can be grouped in two dimensions: objective (availability) and subjective (desirability). The former, quantitatively ranging from rarity to abundance, is subject to collection bias as scientifically defined, while the latter defines the subjective aims of the collector. Both are influenced by cognitive biases, whose discussion is beyond the scope of this short paper, meant just to provide a graphical representation, based on logical Venn diagrams, of those main conditions, and to elucidate the discrete categories at the intersection of the observed sets of items.

**KEYWORDS** Collection, completeness, bias, Venn diagrams

### INTRODUCTION

Any collection, including but not limited to musealized or personal collections of biological specimens (seashells, insects...) can be evaluated more or less complete, by its degree of correspondence with the declared objective of its collector. Any collection develops in time across the phases of inception, growth and possibly completion: at any moment in time, the status of the collection is the expression of the interaction of a very few objective and subjective factors, that can roughly but effectively be grouped in two dimensions, availability and desirability. We will refer to those factors as “conditions”, that must co-exist for a viable collection process.

The graphical depiction of logical sets, in the form of the well-noted Venn diagrams, may help in elucidating the existence of a few discrete categories that influence and define the status of the collection, *e.g.* the number of specimens of a given species, or the number of pictures of a given painter, or the number of wristwatches of a given brand. (Venn, 1880)

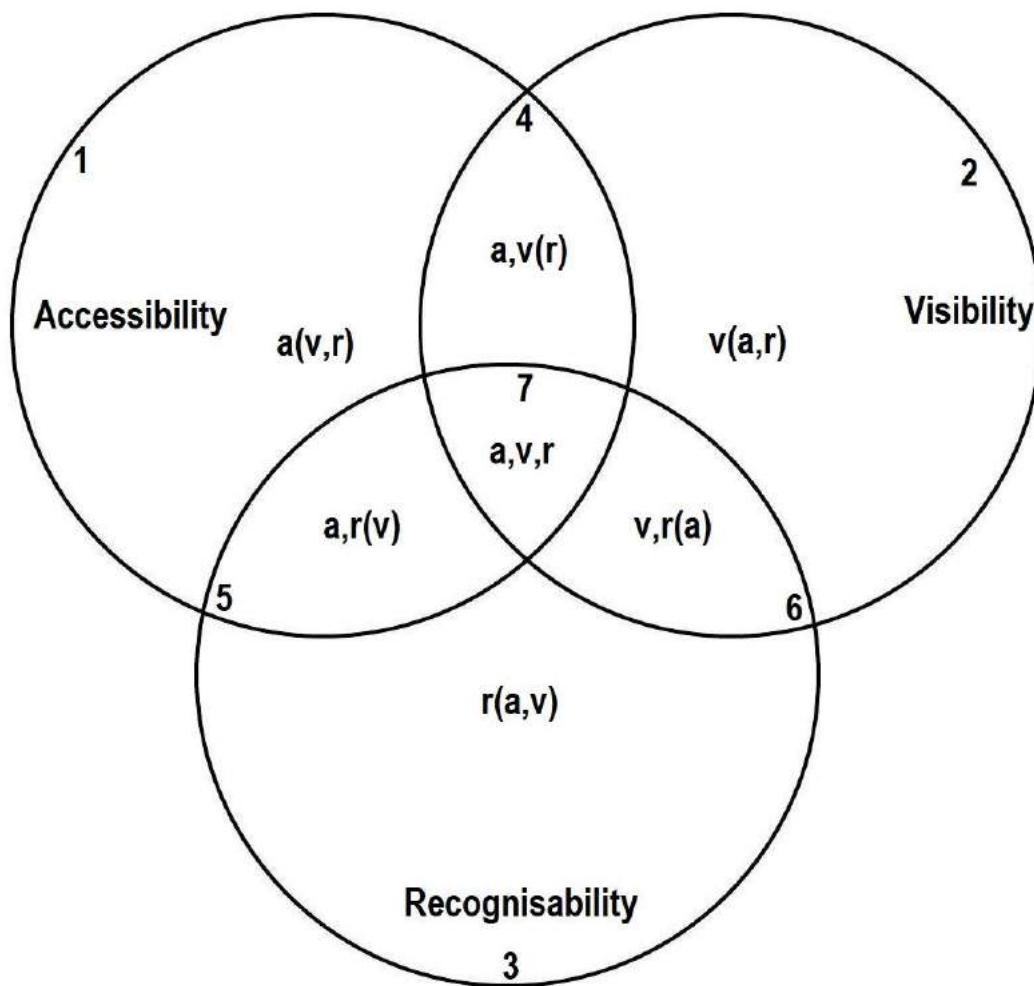
Those graphics will hopefully provide an insight on the reasons why – excluding exceptions – collections usually provide a limited, or biased, representation of the very diversity they are aimed at covering.

## MATERIALS AND METHODS

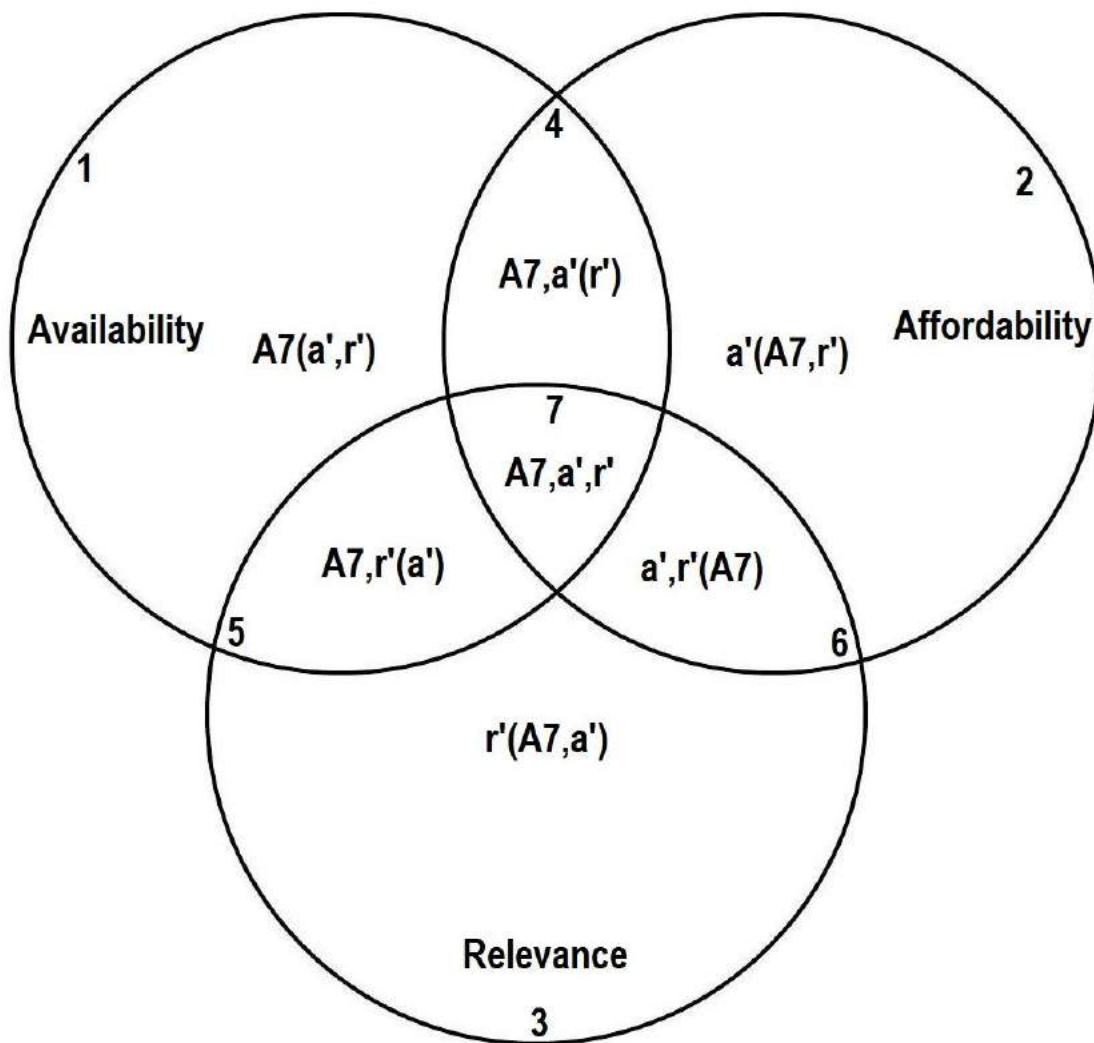
A total of five objective or subjective conditions are considered and traced back to their respective dimension. They include:

- a - Accessibility – an objective condition: to the purpose of this paper, as a factor of availability, accessibility means the physical presence of a given item in the environment or place where it is searched for. An extant item that isn't accessible can't become available. As an example of unavailability, we may cite an historical artifact whose location is unknown.
- v – Visibility (or conspicuousness) – an objective condition, with distinctions: in the context of availability, a visible item is any item that can be located by the means used for its research. For the unaided eye, a very small specimen may be invisible. But if a microscope is used, it may be easily located among several similar or dissimilar items. But many more things hide under the carpet of "visibility", and one question is revealing: to which degree it could reasonably be expected that a collection actually represent its target diversity scope? To answer that question exhaustively, it would be necessary to go very far beyond the aims of this short paper, by delving into the concept of collection bias. An ever-present combination of cognitive biases, such as those discussed by Kahneman (2000), Kuhn (1970) and Plous (1993), collection bias was particularly well explored in the realm of biodiversity studies. Just one reference is included as a tribute to this complex and interdisciplinary issue: as Oliveira et Al. (2016) brilliantly illustrate, cognitive biases are a deceptive and underhanded enemy. Their rich bibliography provides an excellent starter for those interested in collection bias in the natural sciences. Coming to practical terms, the much less ambitious aim of this contribution is exposing and shortly commenting the few most relevant and coarse-grained factors affecting the composition of any collection. In that respect, "visibility", as defined above, may include both the incidence of the collection biases, and their mitigation, based on being aware of their disruptive potential. By remembering that we all see very selectively, under the influence of cognitive biases we can maintain the awareness that one's sight is often much worse than he/she believes.
- r Recognisability – an objective condition, with distinctions: in the context of availability, a recognizable item is any item that displays diagnostic characters, allowing to refer it to a specific entity such as a biological species, or a specific brand and model of wristwatch, or as a painting of a specific painter. Obviously, in actuality the successful recognition depends on the degree of competence of the person in charge of this activity – and this introduces a degree of subjectivity. Furthermore, the degree of contention on nomenclatural / systematics issues related with the item (biological or non-biological) may influence the reliability of the recognition. In the context of biological collections, recognisability brings us into the issue of the quality of scientific literature, potentially opening the Pandora's box of non-peer-reviewed or otherwise unnecessary nomenclatural acts, that cast doubts about the validity of some of the discrete categories on which recognisability is based, that is to say species and subspecies (let alone forms). To the aim of this short paper, let's just consider that the provisions by the International Code of Zoological Nomenclature proved unable to avoid the emergence of questionable taxonomical entities - only the adoption of a voluntary, peer reviewed higher standard, that we may informally define as "serious science", can prevent the appearance of "psychospecies" (Tursch & Greifeneder, 2001) and of pseudo-scientific entities sometime created for commercial purposes only. See also Relevance, more under.
- a' - Affordability – a subjective condition: only the items compatible with the purchasing power of the collector. Furthermore, any collector may modulate his/her expenditure depending on the next condition, relevance. See also Relevance, more under.
- r' - Relevance – a subjective condition, intrinsically linked with the purpose of the collection. The

collection may be purely esthetical or scientific. The same item, or specimen, may be unmissable for a collector, and completely uninteresting for another. Different items may bring different relevance for any collector, with relevance influencing affordability. The main driver of relevance being the recognition of the item (as a given species, brand, model...) it would be very tempting to introduce in the graphics even another factor/condition, reliability. A different course of action was taken: reliability may be viewed as much as an objective (reputation) and a subjective (trust) factor. To the aims of the collecting process, exceptional specimens/items may be collected by the collector himself/herself, by expert researchers or dealers, as well as by occasional researchers or dealers. Even though it's clear that an expert is less prone to recognition errors, nobody would refuse an exceptional item from an inexpert or discredited dealer. It's a responsibility of the collector to ascertain to his/her own satisfaction that the item is desirable. Summarizing, the unclear positioning and conditionality of the "reliability" condition prompted the author to omit it from the graphics, and to put it under the umbrella-definition of "Relevance", as any task delegated to the collector's own judgement.



**Figure 1.** The "Availability" conditions and their interactions.



**Figure 2.** The “Desirability” conditions and their interactions.

Two three-set Venn diagrams are here proposed as Figure 1 and Figure 2 and provide respectively a representation of the logical dimensions availability (A) and desirability (D), including three main factors each.

The areas of intersection are numbered and marked with a notation like  $x(y,z)$  where the unsatisfied conditions are in brackets, the satisfied conditions are before and outside the brackets.

Availability is a precondition for desirability. At the same time, it makes no sense to put the individual factors  $a$ ,  $c$ ,  $r$  as separate sets in Figure 2 overlapping  $c'$  and  $r'$ . For these reasons, in Figure 2 the notation  $A7$  indicates the Availability as a whole, in other words the effectively available items that are available as long as they satisfy the conditions of accessibility, conspicuousness and recognisability (area number 7 in Figure 1).

## DISCUSSION

The reasoning behind Figure 1 and Figure 2 is as follows. First and foremost, to enter a collection, an item must be “available”: at some moment, somebody accessed the item, saw it, recognized it. All the three preconditions are of paramount importance to generate a collectable item. Let’s comment in detail the seven areas of Figure 1:

- Area 1 –  $a(v,r)$ : the item is there where one would expect, yet nobody is specifically looking for it. Even if it was clearly visible, it would not be recognized. This is a suitable definition for most undescribed, small species that usually emerge only with a massive or invasive action of sampling is performed, with no other purpose than “seeing what one gets”.
- Area 2 –  $v(a,r)$ : the item/specimen would be visible, if just someone could access it and recognize it – this area covers the stage set for the “barn finds” of rare cars as well as the stage of any epiphany taking place in other areas of the diagram.
- Area 3 –  $r(a,v)$ : a suitable area for the realm of cryptozoology, as well as for the quest for the Holy Grail. Many species that would be recognizable, if only they were easier to see, are awaiting to be rediscovered, but nobody is there to try!
- Area 4 –  $a,v(r)$ : like Poe’s “Purloined Letter”, rare items lay in plain sight, in the small shop window or among the commonest species in well accessible places. This area would cover any cryptospecies within a well-established group, as well as the rare trove we are unknowingly keeping in our own attic. Similarly, this area may cover the case of many species that could be reached with minimal effort.
- Area 5 –  $a,r(v)$ : field research, or any other systematic search for something, performed without proper instrument or without the right attention, will make us miss the opportunities. as an example, the relevant collection bias against small specimens may transform in rarities even common species with the only distinction of the small size. That way, «difficult to notice» becomes «rare».
- Area 6 –  $v,r(a)$ : It isn’t there! You would easily see it and recognize it but looking for a species outside its range, or looking for an item in the wrong shop, brings one right into this area.
- Area 7 –  $a,v,r$ : all set up: *veni, vidi, vici*. A candidate item / specimen is ready for the collectors. To the aim of this paper, we postulate that whenever the specimen is accessible, visible and recognizable, it will be collected.

As stated before, to candidate an item for the collection all the conditions need to be satisfied: hence, A7 is one of the three condition in Figure 2, that includes the following areas:

- Area 1 –  $A7(a',r')$ : a pricey item indeed, but it’s no problem, as long as the collector is uninterested.
- Area 2 –  $a'(A7,r')$ : even though the item is surely affordable, it’s currently unavailable. This situation causes no concern, because the collector is uninterested.
- Area 3 –  $r'(A7,a')$ : a situation where the collector is eager to acquire an interesting item, but it’s so rare that it’s currently unavailable. Furthermore, it would probably be unaffordable.
- Area 4 –  $A7,a'(r')$ : specimen affordable but uninteresting – may be purchased just because it’s cheap, but iterating this pattern results in a collection where cheaper items are over-represented, regardless of their relevance.
- Area 5 –  $A7,r'(a')$ : a relevant specimen is available, but unaffordable – purchase delayed,

- hoping that wealthier collectors don't harbor the same likings.
- Area 6 –  $a', r'(A7)$ : no relevant and affordable specimen is available – the collection does not increase.
  - Area 7 –  $A7, a', r'$ : the affordable and relevant specimen/item enters the collection under the ideal conditions of availability and desirability.

## CONCLUSIONS

Any collection is a compromise. Its scope may be ambitious but, unless great resources and manpower are available, what one gets is a “fact-based collection”, the fruit of concurring availability and desirability, two dimensions based on multiple and interacting factors. Almost invariably, the commonest species or items will be over represented while even big investments shan't be able to overcome the objective shortage of the most uncommon material.

By exposing in an intuitive way the interacting and often unnoticed and inexplicit factors that govern the evolution of a collection through time, this contribution will hopefully provide a better understanding of the collecting process, or even set the bases for a better purchasing strategy.

## REFERENCES

- Kahneman, D. 2000.** Evaluation by Moments: Past and Future - Chapter 38 in Kahneman D. and Tversky A. (Eds.), *Choices, Values and Frames*, New York: Cambridge University Press.
- Kuhn, T.S. 1970.** *The Structure of Scientific Revolutions*, Chicago: The University of Chicago Press.
- Oliveira, U. et al., 2016.** The strong influence of collection bias on biodiversity knowledge shortfalls of Brazilian terrestrial biodiversity. *Diversity and Distributions* 22:1232–1244.
- Plous, S. 1993.** *The Psychology of Judgment and Decision Making*, McGraw-Hill. ISBN: 9780070504776, 233.
- Tursch, B. and Greifeneder D. 2001.** *Oliva Shells. The genus 'Oliva' and the Species problem*, Recanati (Italy), Tecnostampa.
- Venn, J. 1880.** On the Diagrammatic and mechanical representation of Propositions and reasoning, *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science*, Series 5, 10:59, 1-18.

### Cite as:

**Brizio, C. 2023.** The Completeness of Thematic Collections - a Venn Diagram-Based Conceptual Approach. *The Festivus* 55(2):131-136. <http://doi:10.54173/F552131>