Concept Paper

Benchmarking in Taxonomy: A Commentary on the Role of the Holotype

George Scott *

Concept Paper

Benchmarking in Taxonomy: A Commentary on the Role of the Holotype

George H. Scott

Paleontology, GNS Science, P.O. Box 30368, Lower Hutt, New Zealand; George.Scott@gns.cri.nz

Abstract

Benchmarking in taxonomy is viewed both as establishing a specimen as standard of reference and as a process for optimizing that process. Here it is founded on vision theory that recognition of specimens, as for all objects, is personal to the observer and is based on stored exemplars (benchmark images) in their memory. A special feature of a holotype as a scientific benchmark is that it has been published with a Linnaean name permanently attached. This concept is generalized to an entity that includes all specimens published by subsequent taxonomists with that name attached (a labelled specimen database). As a record of usage it integrates all published images with a Linnaean name. It promotes an inquiry into processes for selection of such specimens. In the conventional model of practice taxonomists categorize specimens using their stored representations of already identified individuals; the process is immediate, acute and autonomous but is largely concealed; a specimen may be selected as a benchmark but its typicality is not revealed. As a remedy, a population model of practice is advocated wherein the basic autonomous visual process is supplemented by objective data about a specimen and the probability of its position within a potential source population.

Keywords: benchmarks; taxonomy; holotype; vision science; referents; biological nomenclature

1. Introduction

In its original usage 200 years ago a 'benchmark' was a surveying term that referred to a mark in the ground which enabled instruments to be re-established in their original location: it was a reliable point of reference, a ground truth location. Although this usage continues, it has evolved in the digital age to include processes seeking to optimize performance (Bartz-Beielstein, 2020, Moriarty & Smallman, 2014) and is widely used in science and technology. A principal application in the biological sciences has been in computational genomics where it refers to optimization of processes for decoding molecular data.

Considered as a point of reference, as in its original use, 'benchmark' is seldom used in Linnaean taxonomy either by practicing taxonomists or by philosophers of science and semantics yet, as identified by ICZN (1999), the holotype specimen appears to be a prime example. Examined here is its use by Brummer & Kucera (2022) in their review of the taxonomy of living planktonic foraminifera. They aim to benchmark species concepts, noting that the taxonomy should be robust and operational. Their usage of 'benchmark' is simple: for species, they confirm that a holotype specimen exists and that the name is not a subjective synonym (ICZN, 1999 art. 61). Otherwise expressed, they assess whether the species name been validly typified. While this is a strict interpretation of a benchmark specimen in Linnaean taxonomy, it overlooks its role in the recognition of such taxa, either as new species or as specimens of already named species. Several aspects of these roles are considered here.

2. Morphospecies: Their Recognition

In cognitive neuroscience 'categorization' (more generally 'classification') denotes a process whereby a group of objects is picked out by an individual's mental representation (concept) of the group. It allows for determination of objects belonging to the group (Medin & Coley, 1998); visual similarity is a primary determinant of membership. Current research indicates that the mental representation may be based on exemplars (members) of the group already stored in visual memory (Smith, 2016; Ashby & Rosedahl, 2017), or on prototypes formed from summed similarities of exemplars. These processes are located separately in the brain but may operate in tandem (Blank & Bayer, 2022). Category learning is aided by the extent of similarity among exemplars and may be assisted by textual data (Hughes & Thomas, 2021). Importantly, all of these processes are personal to each observer.

In practice, building knowledge of a morphospecies begins with a taxonomist recognizing a new (=unnamed, i.e., no imagery recall) group within his collections or in the literature, depositing a specimen (holotype) in a recognized repository, and publishing a Linnaean name and description for it. This makes the species available in the public domain as a named taxon, as Brummer & Kucera (2022) showed. The holotype is the name bearer and as such qualifies as a benchmark, or equivalently as a voucher specimen in a machine learning application. That the species name is bound to that specimen creates a widely discussed problem in its use (Michaelson, 2024) and is a focus of this commentary.

How the author's concept is to be promulgated is not addressed by ICZN (1999) and seldom is in the wider literature, yet is vital for its use as a referent. The following representation seeks to mirror conventional practice for referencing the name of a newly recognized taxon as a visual entity. In it 'labelled' is used for specimens, including images thereof, to which the Linnaean name is attached, say as a label on the specimen or on its mountant, or in captions for illustrations of the specimen. In this sense the holotype is the primary labelled specimen (PLS). Knowledge of the species is incremented and disseminated as other taxonomists view the deposited museum specimen, or its published image, and recognize similar specimens in their collections. They might isolate and label such specimens, subjectively determined (SLS), with that name for their reference or publish illustrations, thereby extending the set of labelled specimens. Some may have inspected the holotype (Figure 1) but commonly this is infeasible. Increasingly, available sources of visual knowledge of species are SLS specimens shown in print or digital publications. They form an essential guide for today's taxonomists to develop their personal concept of a species and also serve as an important adjunct to the limited capacity of our short term memory (Brady et al., 2011).

For a given species, labelled specimens form an incremental record of its identification that expands out from its original publication. Ostensibly, the expansion is an historical record of specimens identified by later taxonomists labelled using the name of the holotype specimen: it is termed a labelled specimen knowledge base (LSKB). For a recently described species individual entries in the LSKB are readily traceable online and the history of usage is viewable. However, as taxonomists seldom record which prior labelled specimens they inspected when forming their concept of a species, for many taxa their LKSB resembles a virtual entity.
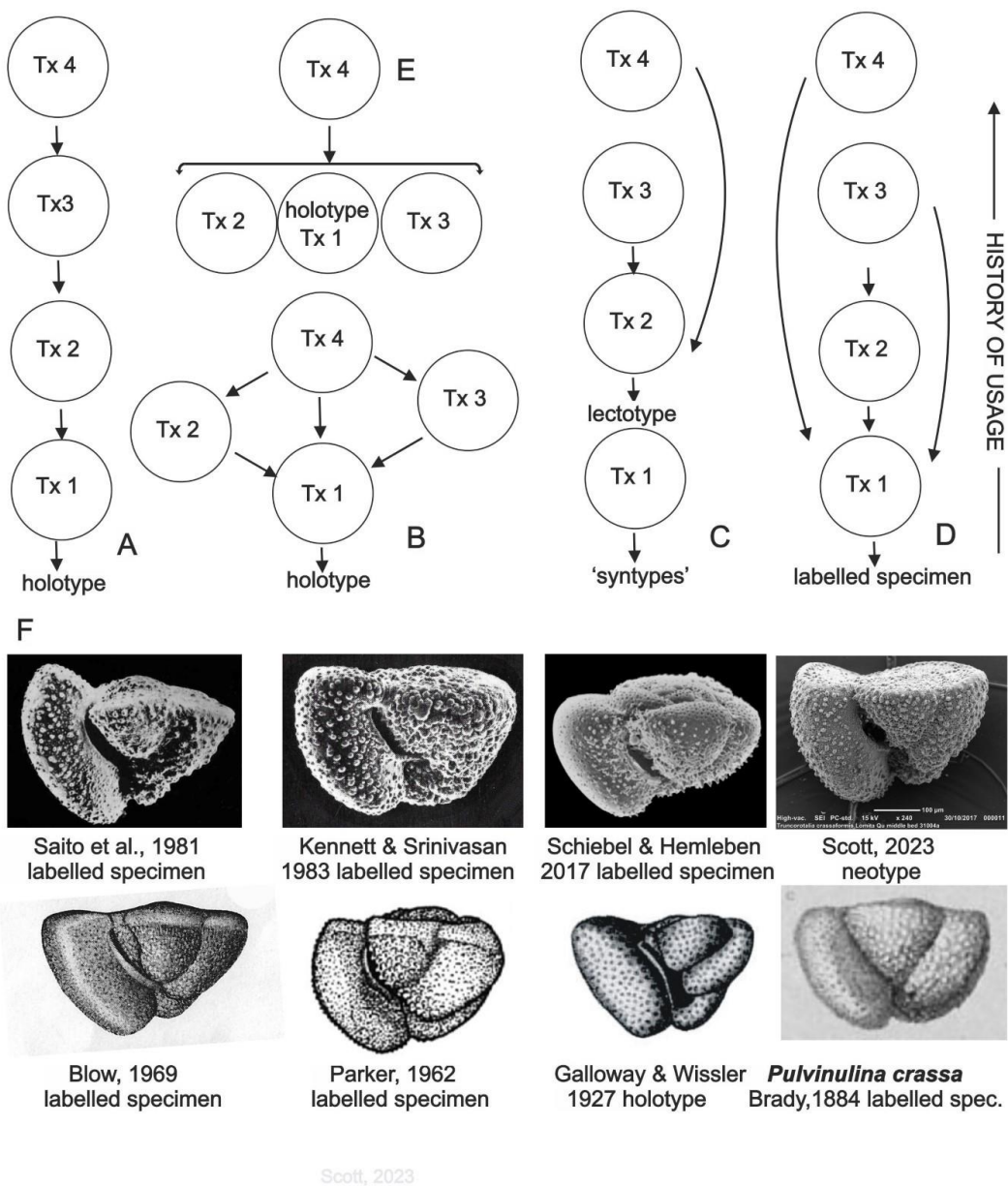
Figure 1. (A-E) Hypothetical LSKB viewed as flow maps of visual information about species. Arrows indicate a previously labelled specimen used as a guide for a taxonomist (Tx n) to identify and label a specimen in his collection. (A) compares with the chain-like structure of the historical theory of causal reference in semantics (Michaelson, 2022) and might map a preference for viewing the last-published image because of improvements in image quality over time. (B) Tx 4 views all available imagery. (C) models some planktonic foraminiferal taxa named in the early 19th century (Banner & Blow, 1960) for which authors isolated several specimens ('syntypes') but did not label a holotype; published images were of models.   (D) shows the significance of scientific atlases (Daston & Galison, 1992) in taxonomy; in this example, as in several real examples, the atlas is not rooted in the holotype, (E) shows the role of online image aggregator sites which are becoming the default sites for viewing labelled specimens. (F) Labelled specimens of   *Globorotalia crassaformis* (Galloway & Wissler, 1927) from selected references.

3. Discussion

3.1. Conventional Model

This model (CM) relies on the operations of our vision system for construction of a LSKB and inherits its defects for recognition of groups based on a primary labelled specimen. Nominally, a

LSKB is an archive of images whose name is a homonym of the holotype: it uses the name of the holotype but references different specimens. This referencing problem is generally ignored by taxonomists (e.g., Braby 2024; Ruedas 2024) but is widely discussed in the philosophy of the natural sciences (e.g., Ereshefsky & Reydon, 2015 and their citations; Michaelson, 2024). Of relevance here is Crane (2021, p. 12195) who regards the holotype as a sample and argues that "One needs a particular explanatory purpose to pick out a particular population by a species baptism. If that is left open, the name may not latch onto any determinate population". Although this is not an operational solution, it recognizes the need to relate a holotype to its source population and is consistent with the following model.

3.2. Population Model

This model (PM) seeks to remedy problems raised by the conventional model. It accepts that 'species' is a group concept for biological specimens recognized as similar tangible objects, a view that compares with Darwin (1861 p. 159) "To sum up, I believe that species come to be tolerably well-defined objects ...". To expand this definition, I follow Simpson (1961, p. 65) that "... Populations, not individuals, are the units of systematics ..." and Beckner's (1959) view that species are polytypic populations in space/time. The problem then for a species proposal is to identify a specimen that to serve as the primary benchmark for that population, or as a benchmark for a local named population. It leads to a quite different, probabilistic, search strategy wherein the opaque operations of the personal perception process of CM are supplemented by strategies that create probabilistic, contestable data about the position of the holotype or other labelled specimen in its source population. The schema of the population model used here follows these steps. 1. The taxonomist uses their visual perception to select specimens they provisionally identify as species A or similar thereto. 2. Morphometric data on a functional trait are gathered. 3. Statistical analyses provide probabilities of the selected specimens belonging to one population. 4. The taxonomist evaluates this statistical grouping against their knowledge of the functional significance of the trait. 5. A benchmark or other voucher specimens are selected, informed by the probability that they belong to Species A.

The trait in the examples of this study is the axial outline of planktonic foraminifera shells. Experimental neuroscience shows that, subliminally, the outline is an important guide to object recognition (Sprote et al, 2016; Elder, 2018). Here it captures most of a specimen's ontogeny and is regarded as serving hydrodynamic and trophic functions. The normal model is preferred for analysis because it provides a suitable measure of central tendency, which might be associated with adaptation towards an optimum form in the shape trait.

The primary focus of PM is to demonstrate the probability that specimens in a sample are drawn from one population (Figure 2). Its value for accessing the relation of benchmark specimens to other populations is shown in Figure 3. Theoretically, PM might be viewed as an optimizing strategy for selection of representative specimens such as benchmarks. Nevertheless, the statistical data remain open to interpretation, biologically: the taxonomist retains discretionary judgment when, for example (Figure 2 #30), a specimen is flagged as unlikely to belong to the population when it is interpreted as a juvenile (Scott, 2025).

Operationally, PM requires that the population from which the holotype was sourced is analysed and that data from that analysis are publicly available to allow them to be plotted in an analysis of a local population that a taxonomist considers to represent the holotype population (e.g., Figure 3). This is a project that organizations like that represented by Brummer & Kucera (2022) might consider.
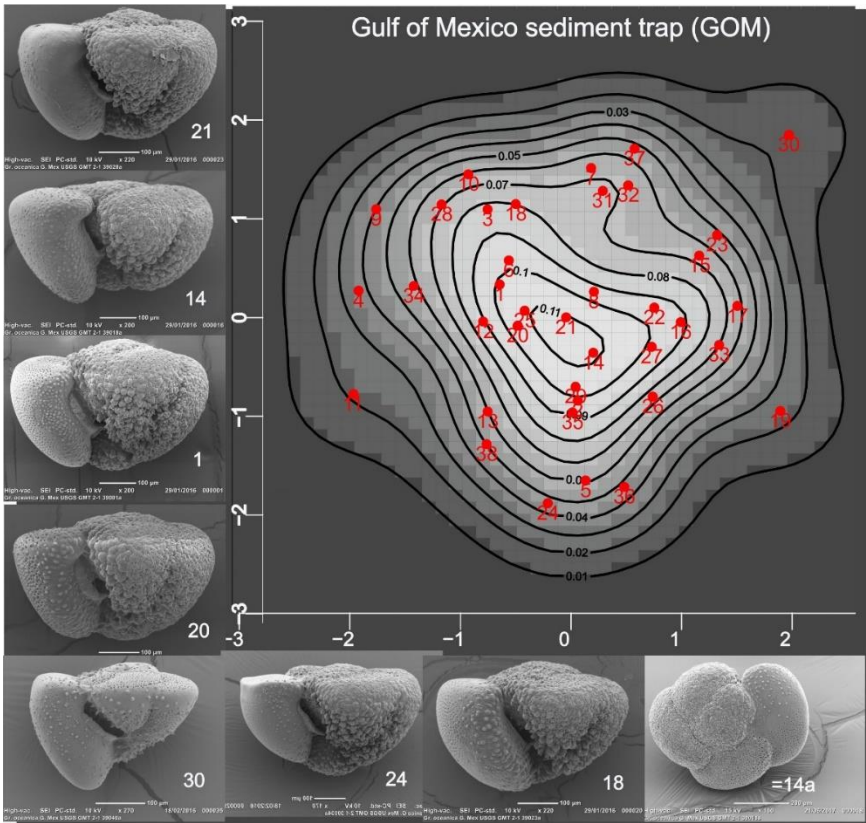
Figure 2. Density map showing probabilities that specimens in the Gulf of Mexico GMT21 sediment trap, a time series (2008–2012) of foraminiferal and particulate flux at 700 m on the northern Gulf of Mexico continental shelf (27.5° N; 90.3° W), belong to a normal (gaussian) shape population. A benchmark specimen might be selected from those with highest probabilities. Refer to Scott (2025) for details.
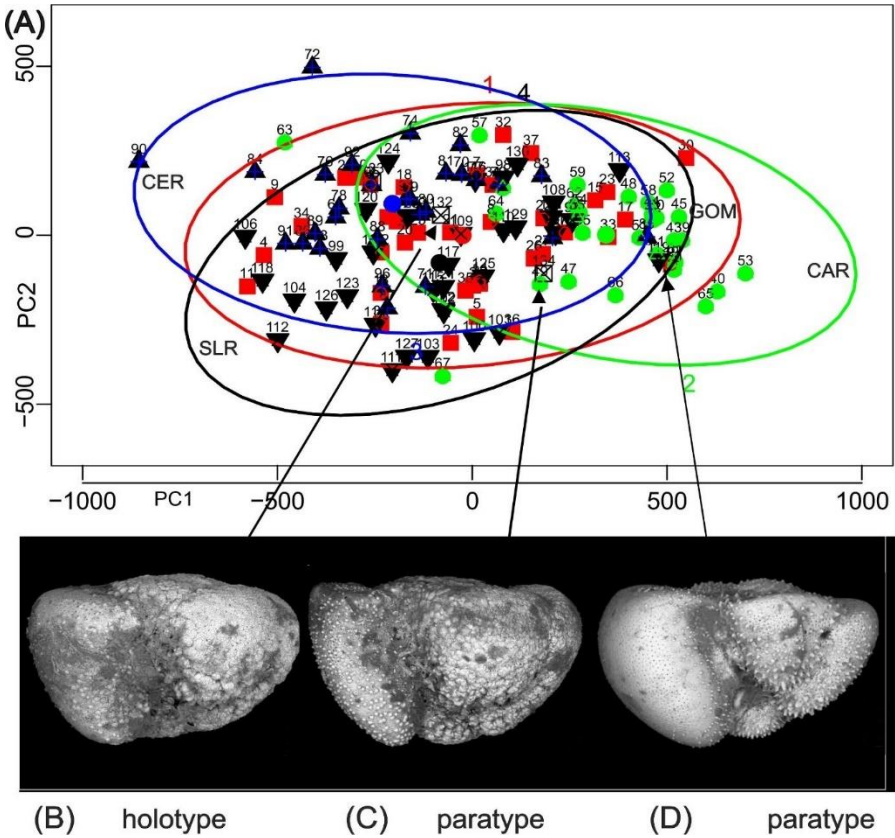
Figure 3. (A) Data ellipses (axial data) for a procrustes analysis of pooled samples Gulf of Mexico (GOM), Cariaco Basin (CAR), Ceara Rise (CER). Sierra Leone Rise (SLR). These collections are from the source region of *Truncorotalia oceanica* Cushman & Bermudez, 1949. (B-D) Holotype and paratypes of *Truncorotalia oceanica* projected on to (A). Refer to Scott (2025) for details.

4. Conclusion

Brummer & Kucera (2022) offer an account of the application of benchmarking modern planktonic taxa that is limited to verifying the existence of a formally named specimen, its subjective synonyms and specification of diagnostic characters, but without imagery thereof. They do not address how an operational taxonomy is to be developed from these data.

ICZN (1999) rules relate only to the formation of a valid name for a species and the curation of one specimen, the holotype or benchmark specimen. They do not provide guidance on usage of the name of that specimen: its role as a referent. In conventional practice taxonomists identify individual specimens using their personal concept of a species, which is informed by their stored mental representations of previously viewed specimens, or their images, to which the name has been applied. Collectively, the named images form a database of referents, which, excepting the holotype, are subjective interpretations of the name and personal to each taxonomist's species concept.

On the workbench, morphospecies comprise tangible specimens primarily grouped by their common visual attributes. Although conventional practice, reliant on the autonomous processes of visual memory, is highly efficient for specimen recognition and grouping of such objects, its mechanisms are cryptic, and outcomes are personal to the viewer. They do not provide objective data about the relation of a labelled specimen to the name bearing specimen (benchmark) and are unsuitable for species when viewed as polytypic populations. Contrarily, for the population model the taxonomic problem is to identify their populations and generate contestable, objective data about specimen relationships to facilitate selection of representative specimens.

Expectedly, because of its history of usage, efficiency and general acceptance, the conventional model will remain normative into the future and accusations of "stamp collecting" (Johnson, 2007) will continue. Acceptance of the population, or similar models, will be contingent on wider appreciation of the constraints imposed on recognition of a natural group when its name is bound to one specimen. Impeding wide usage of the population model for many cohorts is the present limited availability and cost of laboratory equipment for rapid specimen positioning and capture of 2-3D image data.

References

1. Bartz-Beielstein, T.; Doerr, C.; Berg, D. van den; Bossek, J.; Chandrasekaran, S.; Eftimov, T.; Fischbach, A.; Kerschke, P.; Cava, W. L.; Lopez-Ibanez, M.; Malan, K. M.; Moore, J. H.; Naujoks, B.; Orzechowski, P.; Volz, V.; Wagner, M.; Weise, T. Benchmarking in Optimization: Best Practice and Open Issues. arXiv December 16, 2020. https://doi.org/10.48550/arXiv.2007.03488.

2. Beckner, M. *The Biological Way of Thought*; Columbia University Press, 1959. https://doi.org/10.7312/beck93140.

3. Blank, H.; Bayer, J. Functional Imaging Analyses Reveal Prototype and Exemplar Representations in a Perceptual Single-Category Task. *Communications Biology* **2022**, *5* (1), 896. https://doi.org/10.1038/s42003-022-03858-z.

4. Blow, W.H. Late Middle Eocene to Recent Planktonic Foraminiferal Biostratigraphy. In *Proceedings of the First International Conference on Planktonic Microfossils*; Brill: Leiden; Vol. 1, pp 199–422.

5. Braby, M.F.;Hsu, Y-F; Lamas, G. How to Describe a New Species in Zoology and Avoid Mistakes. *Zoological Journal of the Linnean Society* **2024**, *202*, 1–16.

6. Brady, H.B. Report on the Foraminifera Dredged by HMS Challenger during the Years 1873-1876. In *Report of Scientific Results of the Exploration Voyage of HMS Challenger*; Zoology; Challenger Office: Edinburgh, 1884; Vol. 9, pp 1–814.

7. Brady, T,F,, Konkle, T., Alvarez, G.A. A Review of Visual Memory Capacity: Beyond Individual Items and toward Structured Representations. *Journal of Vision* **2011**, *11* (5), 1–34. https://doi.org/doi:10.1167/11.5.4.

8. Brummer, G.-J. A.; Kučera, M. Taxonomic Review of Living Planktonic Foraminifera. *Journal of Micropalaeontology* **2022**, *41* (1), 29–74. https://doi.org/10.5194/jm-41-29-2022.

9. Crane, J.K. Two Approaches to Natural Kinds. *Synthese* **2021**, *1999*, 12177–12198. https://doi.org/10.1007/s11229-021-03328-9.

10. Cushman, J.A., Bermudez, P.J. Some Cuban Species of Globorotalia. *Contributions from the Cushman Laboratory for Foraminiferal Research* **1949**, *25* (2), 26–45.

11. Darwin, C. *Origin of Species*; D. Appleton: New York, 1861.

12. Daston, L.; Galison, P. The Image of Objectivity. *Representations* **1992**, *40*, 81–128. https://doi.org/10.2307/2928741.

13. Elder, J. H. Shape from Contour: Computation and Representation. *Annu Rev Vis Sci* **2018**, *4*, 423–450. https://doi.org/10.1146/annurev-vision-091517-034110.

14. Ereshevsfy, M., Reydon, T.A. Scientific Kinds. *Philosophical Studies* **2015**, *172*, 969–986. https://doi.org/DOI 10.1007/s11098-014-0301-4.

15. Galloway, J.J., Wissler, S.G. Pleistocene Foraminifera from the Lomita Quarry, Palos Verde Hills, California. *Journal of Paleontology* **1927**, *1*, 35–87.

16. Hughes, G.I., Thomas, A.K. Visual Category Learning: Navigating the Intersection of Rules and Similarity. *Psychonomic Bulletin & Review* **2021**, *28*, 711–731. https://doi.org/10.3758/s13423-020-01838-0.

17. *International Code of Zoological Nomenclature: Code International de Nomenclature Zoologique*, 4th ed.; International Commission on Zoological Nomenclature, Ride, W. D. L., International Trust for Zoological Nomenclature, Natural History Museum (London, England), International Union of Biological Sciences, Eds.; International Trust for Zoological Nomenclature, c/o Natural History Museum: London, 1999.

18. Johnson, K. Natural History as Stamp Collecting: A Brief History. *Archives of Natural History* **2007**, *34* (2), 244–258. https://doi.org/DOI: 10.3366/anh.2007.34.2.244.

19. Kennett, J. P.; Srinivasan, M. S. *Neogene Planktonic Foraminifera: A Phylogenetic Atlas*; Hutchinson Ross: Stroudsburg, Pa, 1983.

20. Medin, D.L.. Colely, J.D. Concepts and Categorization. In *Perception and Cognition at Century's End*; Academic Press: San Diego, 1998; pp 403–440.

21. Michaelson, E. The Vagaries of Reference. *Ergo* **2022**, *9* (52), 1433–1448. https://doi.org/10.3998/ergo.3115.

22. Michaelson, E. Reference. In *The stanford Encyclopaedia of Philosophy*; 2024.

23. Moriarty, J. P.; Smallman, C. En Route to a Theory of Benchmarking. *Benchmarking: An International Journal* **2009**, *16* (4), 484–503. https://doi.org/10.1108/14635770910972423.

24. Parker, F.L. Planktonic Foraminiferal Species in Pacific Sediments. *Micropaleontology* **1962**, *8*, 219–254.

25. Ruedas, L.A., Norris, R.W., Timm, R.M. Best Practices for the Naming of Species. *Journal of Mammology* **2024**, *XX*, 1–9.

26. Saito, T., Thompson, P.R., Breger,D. *Systematic Index of Recent and Pleistocene Planktonic Foraminifera. 190 Pp., 56 Plates. Tokyo: University of Tokyo Press. Price £15.00. ISBN 0 86008 280 6.*; University of Toyko: Toyko, 1981.

27. Schiebel, R., Hemleben, C. *Planktic Foraminifers in the Modern Ocean*; Springer: Berlin, 2017.

28. Scott, G. H. A Replacement Neotype for Globigerina Crassaformis Galloway & Wissler, 1927. *Journal of Foraminiferal Research* **2023**, *53* (4), 397–402.

29. Scott, G. H. An Introduction to Globorotalia (Turborotalia) Oceanica Cushman & Bermudez, 1949 in the Modern Ocean. **2025**. https://www.preprints.org/manuscript/202502.1462/v1.

30. Simpson, G. G. *Principles of Animal Taxonomy*; Columbia University Press, 1961. https://doi.org/10.7312/simp92414.

31. Smith, J. D. Prototypes, Exemplars, and the Natural History of Categorization. *Psychon Bull Rev* **2014**, *21* (2), 312–331. https://doi.org/10.3758/s13423-013-0506-0.

32. Spröte, P.; Schmidt, F.; Fleming, R. W. Visual Perception of Shape Altered by Inferred Causal History. *Sci Rep* **2016**, *6* (1), 36245. https://doi.org/10.1038/srep36245.