

KNO-010 · Classification as a Knowledge Tool

KNO · Ways of Knowing

Juan Segurola

v0.1 · EN

All Rights Reserved

Article Abstract

A classification is a working instrument, not a picture of the world. It is built so that something further can be done: a specimen found again, a death counted, a book shelved, a population governed. This article reads classification as a tool with declared purposes and measurable costs. No working system Bowker and Star examined meets the three properties an ideal one would have, the ICD shows how one scheme handles what its categories miss, and grouping, naming and identifying are separate operations a single scheme quietly fuses. The species case shows a category that is indispensable and unsettled at once; rank shows a convention that is not comparable across groups. Folk classification sits beside scientific taxonomy rather than below it. Applied to people, the same properties acquire consequences a specimen does not suffer. Revision is paid for twice, and part of the bill falls on everything already counted.

In 60 Seconds

Classification looks like description and is not. A scheme is a set of boxes built so that some further operation can run, and its categories were shaped by that operation. Bowker and Star name three properties an ideal classification would have and report that none of the working systems they examined satisfies all three. A scheme that is not complete has to place what its categories did not anticipate, and where a residue sits tells a reader how that scheme handles it. Three operations hide inside the one word: grouping, naming and identifying, which fail differently and are repaired differently. Biology supplies the sharpest case, a species category that is used everywhere and defined by no agreed rule, and a rank system whose levels are not comparable between beetles and primates. Folk schemes are not earlier drafts of scientific ones; they answer a different question. When the things sorted are people, the ordinary properties acquire consequences, and two documented cases show a library scheme carrying its period's assumptions forward and a state scheme deciding what a person was. Revision costs money now and comparability afterwards.

Why This Question Matters

Classifications are not consulted, they are inhabited. A diagnosis code decides what a death is counted as, a shelf mark decides what a reader finds, a population category decides what a state does to a person. Because the categories were shaped by the operation they had to support, a scheme read as a report on how things are divided lends the world a structure that came from the work. That misreading is cheap to make and expensive to hold: it turns a revisable convention into a fact, and it hides the decision that a category records.

1. What a classification is asked to do

A classification is not a description. It is a device for putting things into groups so that further work becomes possible, and its groups are made by the work they will have to support. Bowker and Star define a classification as a segmentation of the world and a classification system as "a set of boxes (metaphorical or literal) into which things can be put to then do some kind of work" [1]. The phrase to note is the last one. A scheme that supports no further operation is not a poor classification. It is not one.

This immediately makes classification an engineering problem with an epistemic surface. The engineer asks whether the scheme lets the operation run. The reader of the scheme asks what the operation required the world to look like.

Bowker and Star set out three properties an ideal classification would have: it operates by consistent and unique classificatory principles, its categories are mutually exclusive, and the system is complete [2]. Each is reasonable on its own. Bowker and Star report that none of the working systems they examined satisfies all three, and they extend that beyond their sample as an argued doubt rather than a proof: "No real-world working classification system that we have looked at meets these 'simple' requirements and we doubt that any ever could" [3].

This is the first thing to understand about classification and it is not a complaint. The three requirements are a standard against which real schemes are read, not a specification the systems examined were found to meet. A scheme that fails them is normal. A scheme presented as meeting them is being misdescribed.

Where That Leaves Us

A classification is a device built so that a further operation can run, so its groups were shaped by that operation rather than found in the world. Three properties would make one ideal: consistent principles, mutually exclusive categories, completeness. None of the systems Bowker and Star examined satisfies all three, and they doubt any could. Failing that standard is ordinary; claiming to meet it is the thing to distrust.

2. The residue

Incompleteness is the requirement whose failure is easiest to see, because a scheme that is not complete has to do something with whatever its categories did not anticipate. The ICD shows one explicit way a working classification handles cases its main categories do not capture.

In the International Classification of Diseases the residue is explicit and it is distributed. Bowker and Star observe that the phrase "not elsewhere classified" appears throughout the entire ICD but nowhere as a top-level category [4]. That distribution is informative. A single top-level box for the unclassified would gather the unnamed into one place. Distributing the residue through the branches instead keeps the effects of uncertainty local to the branch in which they arise. That is what the observation establishes; what a reader concludes from the arrangement about the scheme as a whole is not established by it.

A residual category is not a gap in a classification. It is a working part of it, and reading a scheme well means reading where its residues sit and how much traffic they carry.

Where That Leaves Us

Incompleteness is the requirement whose failure shows most plainly, because whatever the categories did not anticipate still has to be placed. In the ICD the residual phrase runs through the branches and never sits at the top, which keeps uncertainty local to wherever it arose. That distribution is what the observation establishes. Reading a scheme well means noticing where its residues sit and how much traffic passes through them.

3. Three operations that a single scheme hides

Classification is regularly used as one word for three different operations, and separating them removes a large amount of confusion.

Grouping decides which things belong together. It is the operation the previous sections describe.

Naming attaches a label to a group. It is governed by conventions of its own, and those conventions can be changed without changing a single grouping. Nothing about the world decides whether a group is called by one name or another; a code of nomenclature does.

Identifying decides which group a particular thing belongs to. It runs on evidence available at the moment of the decision, which is often much less than the evidence that established the group.

The three come apart in practice. Two workers can agree on the groups, disagree on the names, and both be right about the code they follow. Two workers can agree on the names and place the same specimen differently because the identification key resolves what they can see rather than what the group is defined by. **When a scheme is criticised, it is worth asking which of the three operations is actually being criticised**, because the remedies are unrelated.

Where That Leaves Us

One word covers three operations. Grouping decides what belongs together, naming attaches a label under a code that can change without moving anything, and identifying places a particular thing using whatever evidence is at hand. They come apart in practice: agreement about groups survives disagreement about names. Sorting a criticism into the three is worth the trouble, because their remedies have nothing to do with each other.

4. The species problem

Biology gives the clearest case of a classification whose most used category has no agreed definition. Wells and colleagues state that species "remain controversial in terms of definition, delimitation and reality" [5]. This is not a minority position and not a recent one. Species are counted, protected by law, listed as extinct and used as the unit of a very large part of biological research, and the unit itself is contested.

Their proposal is worth stating precisely because it is not a retreat. They argue that "treating species as a heuristic can be consistent with a theoretical definition of what species are" [6], the theoretical content being that species are clusters of closely related individuals responding in a similar manner to comparable sets of evolutionary and ecological forces [6]. A heuristic in this sense is not an admission of failure. It is a statement about what kind of instrument the category is: one that does reliable work across a wide range of cases without being reducible to a single decision rule.

The species case shows something general. A category can be indispensable and unsettled at the same time, and the indispensability is not evidence that the definition has been found.

Where That Leaves Us

The species is the unit of much of biology and it is contested, while that same unit is counted, protected by law and listed as extinct. Wells and colleagues call it controversial in definition, delimitation and reality, and argue that treating the category as a heuristic can sit alongside a theoretical account of what species are. Needing a category constantly is no evidence that its definition has been found.

5. Rank is a convention

Above the species the difficulty changes shape. The question is no longer what the category is but what its level means.

Lee and Skinner state that "there are no generally accepted criteria for specifying the rank to which a taxon should be referred" and that difficulties attach to the criteria that have been proposed; they add that "taxa assigned to a particular rank are rarely equivalent entities" across different groups [7]. A family of beetles and a family of primates are not comparable quantities. They share a word.

Phylogenetic nomenclature responded by proposing to name clades without ranks. Benton reports that most proponents hold that ranks should be abandoned, while noting that the PhyloCode itself does not demand this, and he objects that ranks carry usable information about the relative inclusiveness of clades and serve as practical proxies in macroevolutionary and biodiversity work: "why abandon ranks when nothing other than semantic purity is gained by doing so?" [8].

Neither position disputes that ranks are not comparable across groups. What Lee and Skinner report and what Benton argues differ over whether a convention known not to be comparable is still worth keeping for the work it does. They are two stated positions among others, not the two halves of the field. That is a question about the tool, not about nature, and it is the right shape of question to ask of any classification.

Where That Leaves Us

Above the species the question changes from what the category is to what its level means. No criterion for assigning rank is generally accepted, Lee and Skinner report, and taxa at one rank are rarely equivalent across groups. Benton does not dispute that and argues the convention still earns its keep. Their disagreement is about the tool, which is the right kind of question to put to any scheme.

6. Folk classification, placed beside and not below

Scientific taxonomy is one classification of living things among many, and the others are not earlier drafts of it.

Ethnobiological work has found recurrent structure. Atran reports that Berlin established a standard terminology of folk-biological ranks: the folk kingdom rank, the life form rank, the generic or generic species rank, the folk specific rank and the folk varietal rank [9]. The recurrence is the finding. It says that human groups working independently arrive at hierarchies of broadly similar depth, and that one level in particular carries most of the naming and learning.

The fit with scientific taxa is partial and it is uneven in a specific way. Atran states that generic species "often correspond to scientific genera or species, at least for those organisms that humans most readily perceive, such as large vertebrates and flowering plants", and that on occasion they correspond instead to local fragments of biological families, orders, or with invertebrates even

higher taxa [10]. Atran reads that unevenness as tracking perceptual salience rather than phylogeny, and the reading is his contribution to a debate rather than a settled result. It is offered here as the pattern he reports across the cases he examined, not as a rule that fixes in advance where any two schemes will converge.

Neither the recurrence nor the mismatch supports a developmental ladder. A classification that groups by what an organism does, where it is found, or what may be done with it is answering a different question from one that groups by descent, and answering a different question is not answering the same question badly.

Where That Leaves Us

Independent human groups converge on hierarchies of comparable depth, and Atran reports the standard terminology Berlin established for their ranks. The correspondence with scientific taxa is partial and uneven, and Atran's reading of the unevenness is his contribution to a debate rather than a settled result. A scheme that groups by use or habitat answers a different question from one that groups by descent.

7. When the thing classified is a person

Everything above applies to specimens. When the objects are people the same structural features acquire consequences that a specimen does not suffer.

Bowker and Star's term for the central mechanism is torque. Where an individual life and an institutional classification do not align, the authors describe a pressure that runs in both directions: "We call this agglomeration torque, a twisting of time lines that pull at each other, and bend or twist both patient biography and the process of metrication" [11]. Both the biography and the process of measurement are twisted by the encounter, on the authors' own wording, and what is recorded afterwards carries that.

Two documented cases show the range.

Library classification. Thornton traces how ideas about human classification from antiquity informed eighteenth and nineteenth century ideas of race and ethnicity, how those ideas became embedded in the Dewey Decimal Classification, and how the scheme's treatment of these subjects changed across four editions over 130 years, reflecting the presentation of changing dominant societal views [12]. The scheme did not merely contain the assumptions of its period. It carried them forward, edition by edition, into the arrangement of collections built long after the assumptions were abandoned.

State classification of population. Posel's study of twentieth century South Africa describes classification criteria that combined appearance and general acceptance with "habits, education and speech and deportment", and finds that officials "had considerable latitude to interpret the bureaucratic criteria for racial classification in ways that allowed them to draw widely and idiosyncratically on popular litanies of biological stereotypes" [13]. The point for this article is structural and it is severe. A classification that cannot be applied consistently does not therefore fail to operate. It operates through the discretion of whoever applies it, and the resulting record shows a category where in fact there was a decision.

Neither is a misuse of a neutral instrument, and the two are not the same case. A library scheme carried the assumptions of its period forward into the arrangement of collections; a state scheme decided what a person was for administrative purposes. Both show the ordinary properties described in sections 1 to 3 operating on objects that can be harmed by being sorted, and they differ in what the sorting did.

Where That Leaves Us

With people the ordinary properties acquire consequences a specimen never suffers. Torque is Bowker and Star's name for a pressure running both ways, bending the biography and the process of measurement together. A library scheme carried its period's assumptions forward into collections built long afterwards; a state scheme decided what a person was for administrative purposes. Neither shows a neutral tool being abused, and what the sorting did differs.

8. Revision, and who pays for it

A classification that is used for long enough must be revised. Revision is paid for twice: in the present, through implementation and the retraining of everyone who codes, and retrospectively, through the data already collected.

The International Classification of Diseases is the clearest instrument for seeing this because it is revised deliberately, at long intervals, by an international body. Its eleventh revision was adopted by the Seventy-second World Health Assembly on 28 May 2019, in a resolution that "ADOPTS the eleventh revision of the International Statistical Classification of Diseases and Related Health Problems (ICD-11), to come into effect on 1 January 2022, subject to transitional arrangements" [14].

The cost of such a change is measurable. Janssen and Kunst examined cause-specific mortality in six European countries from 1950 to 1999 and found that "despite using a concordance table based on three-digit ICD codes, in 10.8% of the revisions to codes for cause-specific mortality, significant discontinuities, regarded as being due to ICD revisions, remained", concluding that discontinuities arising from coding changes "can lead to substantial changes in long-term trends in cause-specific mortality" [15].

Read carefully, that is a statement about the limits of translation between schemes. Across the revisions those six countries passed through between 1950 and 1999, a concordance table was applied as the best available repair and roughly one revision in nine still left a discontinuity that the authors attributed to the revision itself rather than to any change in health. The figure belongs to those transitions and is not a rate for revisions in general. **A long time series built across revisions of a classification therefore contains two superimposed signals, one about the world and one about the scheme**, and only the first is usually what the series is being read for.

Where That Leaves Us

Revision is paid for twice: now, through implementation and retraining, and backwards, through everything already recorded. Janssen and Kunst measured the second bill in six countries across five decades, where a concordance table repaired what it could and about one revision in nine still carried a discontinuity the authors put down to the coding change rather than to health. The figure is theirs and not a general rate.

9. Reading a classification well

What the preceding sections support is not scepticism about classification. It is a short set of questions to put to any scheme in use.

What operation was it built to support? The categories were shaped by that operation and will fit others less well.

Where is the residue and how much passes through it? A large or growing residual category is the scheme reporting its own strain.

Which of grouping, naming and identifying is at issue? The three fail differently and are repaired differently.

Is the category settled? Indispensable and defined are different properties, as the species case shows.

Do the levels mean anything across branches? Frequently they do not, and the shared word conceals it.

Who or what is being sorted, and what happens to something that does not fit? With specimens the answer is a residual category. With people it can be torque.

What did the last revision cost, and to whom? The bill is paid by everything already recorded.

A classification is a good tool in the ordinary sense: it does specific work well, it deforms what it is not built for, and it has to be maintained. Treating it as a report on how the world is divided attributes to the world a structure that came from the work.

Where That Leaves Us

The article ends where a reader can use it: a short list of questions to put to any scheme in use, about the operation it was built for, its residue, which of the three operations is at issue, whether the category is settled, whether levels mean anything across branches, what happens to what does not fit, and what the last revision cost. What is established, and how, follows.

What We Know, and How

Bowker and Star define a classification system as a set of boxes into which things can be put to then do some kind of work, set out three properties an ideal classification would have, and report that none of the working systems they examined satisfies all three. They observe that the residual phrase runs throughout the ICD but never as a top-level category, and they name torque for the mutual twisting of a biography and a process of metrickation. Wells and colleagues report that species remain controversial in definition, delimitation and reality, and propose treating them as a heuristic. Lee and Skinner report no generally accepted criteria for rank; Benton argues that ranks still do usable work. Atran reports Berlin's folk ranks and their partial correspondence with scientific taxa. Thornton traces the Dewey treatment across four editions and 130 years, Posel documents official latitude in South African racial classification, the Seventy-second World Health Assembly adopted ICD-11 in 2019, and Janssen and Kunst found discontinuities remaining in 10.8 per cent of code revisions across six countries.

What We Still Don't Know

What a species is, since the category is used everywhere and its definition is contested, and the heuristic proposal is a way of working rather than a settlement. Whether ranks should be kept, which Lee and Skinner and Benton answer differently while agreeing that ranks are not comparable across groups. How far folk generic species and scientific taxa correspond, which Atran reads as perceptual salience and offers as his contribution to a debate. And how much of any long mortality series is the scheme rather than the world, since the 10.8 per cent belongs to the transitions studied and is not a rate for revisions in general.

What Could Change Our Minds

An operational criterion for rank that held across groups would turn Benton's practical defence into a stronger one, and its continued absence is what Lee and Skinner report. Ethnobiological work on groups whose environments differ sharply from those already studied would test whether the recurrence Atran describes is as general as it looks. Concordance work that separated coding discontinuity from real change, revision by revision rather than in aggregate, would say how much of a long series is the scheme. And a classification that documented its own residual traffic would let a reader see the strain instead of inferring it.

If You Remember Only One Thing

A classification is built to make an operation possible, so its categories cannot be read as a neutral picture of the world.

Related Articles

- KNO-009 · *Archives, Libraries and Collective Memory*

References

The bracketed numbers in the body are claim numbers, not reference numbers. Each resolves to a row of the claims ledger, which names the source below that supports it.

Atran, Scott. 1998. «Folk biology and the anthropology of science: Cognitive universals and cultural particulars». *Behavioral and Brain Sciences* 21 (4): 547–569. URL estable: <https://philpapers.org/rec/ATRFAT>

Benton, Michael J. 2007. «The Phylocode: Beating a dead horse?». *Acta Palaeontologica Polonica* 52 (3): 651–655. URL estable: <https://www.app.pan.pl/archive/published/app52/app52-651.pdf>

Bowker, Geoffrey C., and Susan Leigh Star. 1999. *Sorting things out: classification and its consequences*. Cambridge, Mass.: MIT Press. ISBN 0-262-02461-6.

Janssen, Fanny, and Anton E. Kunst. 2004. «ICD coding changes and discontinuities in trends in cause-specific mortality in six European countries, 1950-99». *Bulletin of the World Health Organization* 82 (12): 904–913. **Sin DOI localizado**. URL estable: <https://scielosp.org/article/bwho/2004.v82n12/904-913/>

Lee, Michael S.Y., and Adam Skinner. 2007. «Stability, ranks, and the PhyloCode». *Acta Palaeontologica Polonica* 52 (3): 643–650. URL estable: <https://www.app.pan.pl/archive/published/app52/app52-643.pdf>

Posel, Deborah. 2001. «Race as Common Sense: Racial Classification in Twentieth-Century South Africa». *African Studies Review* 44 (2): 87–113. DOI 10.2307/525576.

Thornton, Lisa. 2024. «How Did We Get Here? Race and Ethnicity in Dewey Decimal Classification». *Knowledge Organization* 51 (6): 414–424. DOI 10.5771/0943-7444-2024-6-414.

Wells, Tom, Tom Carruthers, Pablo Muñoz-Rodríguez, Alex Sumadijaya, John R I Wood, and Robert W Scotland. 2022. «Species as a Heuristic: Reconciling Theory and Practice». *Systematic Biology* 71 (5): 1233–1243. DOI 10.1093/sysbio/syab087.

World Health Organization. 2019. *Eleventh revision of the International Classification of Diseases*. Resolución WHA72.15, punto 12.7 del orden del día, 72.^a Asamblea Mundial de la Salud, 28 de mayo de 2019. URL estable: https://cdn.who.int/media/docs/default-source/classification/icd/icd11/a72_29-en_icd-11-adoption.pdf