Measurement

Gary Hardegree
Department of Philosophy
University of Massachusetts
Amherst, MA 01003

1. Introduction 1
2. The Word 'Measure' 2
3. Definitions of Measurement 4
    1. American Heritage Dictionary 4
    2. Dictionary of Consciousness 4
    3. Free On-Line Dictionary of Computing 4
    4. Web Dictionary of Cybernetics and Systems 4
    5. R. McCleary (Web page at UCal-Irvine for course on Research Design) 4
4. An Initial Account of Measurement 5
5. What is a Comparison? 5
6. Is Every Measurement a Comparison? 7
7. Is Every Comparison a Measurement? 7
8. Assessment and Measurement 9
9. Measurement as Classification 10
10. Scales of Measurement 11
11. Nominal Scales 11
12. Ordinal Scales 12
13. Quick Review 14
14. Comparing Differences 14
15. Qualitative Difference-Comparison 15
16. Quantitative Difference-Comparison – Interval Scales 15
17. Standard Examples of Interval Scales 18
18. Ratio Scales 18
19. Quasi-Ordinal and Quasi-Interval Scales 18
20. Absolute Scales 20
21. Differential-Comparison Revisited 20
22. Hierarchical Classification Systems 22

1. Introduction

Probably the most widespread concept in the empirical (observational) sciences is the concept of *measurement*. Indeed, it is generally believed that measurement is a *sine qua non*[1] for the empirical (observational) sciences, and that the rational (deductive) sciences are essentially concerned with the describing the abstract features of measurement. This has a simple-minded version and a more sophisticated version. At the simple-minded extreme, we picture science as delivering numbers, and we picture mathematics as examining numbers in isolation from their origins. On the other hand, the more sophisticated picture is as follows. On the one hand, the *empirical sciences* seek to discover the *patterns* (regularities) in the world, including the natural world and the artificial world(s). On the other hand, the *rational sciences* (most prominently mathematics, less prominently philosophy) seek to discover what patterns there *might be*, irrespective of whether they are instantiated in the actual world. The two

[1] The phrase 'sine qua non' – literally "without which not" – means an essential element or condition.

components work hand-in-hand, neither being more fundamental than the other. We discover patterns, which we generalize abstractly, which in turn suggests other patterns to look for in nature. When we are looking for patterns or regularities, we need to have a prior idea what a pattern might be.

2. The Word 'Measure'

What is measurement? What is a measure? What is it to measure? In answering these questions, we begin by looking at ordinary usage, in order to appreciate the conceptual task ahead of us. First, the word 'measure' comes from the Latin word *m¶ns¿ra* [measure], which gives rise to the cognate[2] word 'mensuration', which is a fancy synonym for 'measurement'. In this connection, notice that this Latin root also appears in 'incommensurable', which means 'not having a common measure'. Also notice this root in the words 'dimension' and 'immense'.[3]

We next notice that the *American Heritage Dictionary of the English Language* (*AHD*) contains over 30 entries for the word 'measure', which testifies to how varied the usage of this word is! Fortunately, these entries seem to be more or less conceptually inter-related[4] – at least those entries in which 'measure' functions as a verb. There are clear outliers among the noun-uses, most prominently the following.[5]

> desperate times call for desperate *measures*
> the congress passed a *measure* to pay for pre-school education

These uses correspond respectively to the following definitions in *AHD.*

14. An action taken as a means to an end; an expedient.
15. A legislative bill or enactment.

These two uses don't seem to have anything to do with measurement; in particular, there does not appear to be a corresponding verb 'to measure'. We would not ordinarily describe taking measures as "measuring", nor would we ordinarily describe passing measures as "measuring".[6]

---

[2] The word 'cognate' literally means having common birth; in Linguistics, it means having a common origin; more generally, it means being closely related.

[3] Even the word 'semester', which denotes a measure of time, is thought to derive from *m¶ns¿ra*.

[4] This is not always the case. Sometimes words enter a language through a variety of sources, so that a given word can end up having different uses that have nothing "really" in common. An especially noteworthy historical example is the word 'algebra', which in Medieval Europe referred on the one hand to the mathematical technique for solving arithmetic equations, and on the other hand to bone-setting! The story of how this happened is itself quite interesting. In AD 825, al-Khowarizmi published the ground-breaking book *Hidab al-jabr wal-muqubala*. This represents the beginning of algebra, where the word traces to the second word of the title – 'al-jabr'. This Arabic word in turn means 'reunification', which seemingly has nothing to do with solving equations. This word also entered Europe about the same time, when the Moors invaded Spain in the eight Century AD. The word 'al-jabr' was adopted in connection with bone-setting (re-unifying), and bonesetters were in fact called 'algebristas'!

[5] I think that definition 15 derives in an obvious way from definition 14. On the other hand, the usage described by 14 only appears in the plural; we can take *measures* to protect ourselves from disease, but we cannot take *a measure* to protect ourselves. On the other hand, the congress can pass one measure or several measures.

[6] On the other hand, it is a well-known phenomenon of English that nearly any noun can be made into a verb. I describe this phenomenon by the following principle: you can *verb* any word in English. Many people rail against these evident barbarisms, but they continue nonetheless. There was a time, not long ago, when the words 'interface' and 'impact' were only nouns. Thus, there may come a time when the congress is described as measuring, when all it is doing is passing measures.

There are two further uses of ‘measure’ that seem a bit out of the mainstream, although they clearly have something to do with measurement broadly understood. These are given by the following definitions in *AHD*.

16. Poetic meter.
17. Music. The metric unit between two bars on the staff; a bar.

As with 14 and 15 above, these uses of ‘measure’ do not have corresponding verbs. On the other hand, both of these concepts are related, indirectly at least, to verbal uses of ‘measure’.

For example, poems are usually divided into verses (lines), each of which is divided into a number of metrical "feet".[7] Individual feet are *classified* according to rhythm, which is *measured* according to the *pattern* of stressed and unstressed syllables. The *meter* (and hence *measure*) of a verse is then officially defined to be the number and manner of the metrical feet in the line. For example, the most common meter in English poetry is *iambic pentameter*, which means that each verse consists of five iambic feet.[8]

Music is also rhythmically organized, the units being called *measures* (also *bars*), and the *meter* of a measure indicates how that measure is rhythmically organized. For example, a 4/4 (four-four) measure consists of four quarter-notes, whereas a 6/8 (six-eight) measure consists of six eighth-notes. In terms of rhythmic stress, the first beat is stressed, and the remaining beats are unstressed.[9]

We have seen that not every nominal use of ‘measure’ gives rise to a corresponding verbal use. The converse phenomenon also arises; not every verbal use of ‘measure’ gives rise to a corresponding nominal use. The most conspicuous examples are in the following.

the revolutionary tribunal *measured* out harsh justice
the judge *measured* every word carefully

These are inter-related, being given the following definitions in *AHD*.

5b. To allot or distribute as if by measuring; mete;
7. To consider or choose with care; weigh:

As indicated in 5b, a closely related word is ‘mete’ which is a largely archaic synonym for ‘measure’. On the other hand, the word ‘mete’ gives rise to the morphemes ‘meter’, ‘metry’, and ‘metric’ which are widely used.[10] The following are common examples.[11]

[7] More properly, poetry that is *rhythmically regular*. Not all poetry can be measured in this way.
[8] For example, an *iambic* foot (an *iamb*) consists of an unstressed syllable followed by a stressed syllable (as in the word ‘delay’). Other examples of rhythmic feet include – *trochee*: ‘season’; *spondee*: ‘horsefly’; *dactyl*: ‘flattery’; *anapest*: ‘seventeen’; *amphibrach*: ‘remember’.
[9] On the other hand, in rock and roll music, which is officially 4/4, there is usually an additional smaller stress on the third beat. So rock music rhythm is a bit of a hybrid between 4/4 and 2/2 (marches and polkas).
[10] The stem ‘metro’ is also used occasionally – as in ‘metronome’ and ‘hypermetropia’, which refers to an eye malfunction usually called ‘farsightedness’. Note, however, that ‘metropolis’ and ‘metropolitan’ are *false cognates*. Curiously, in these words, the prefix derives, not from *metron*, but from *m¶t¶r*, which means ‘mother’; thus, a metropolis is a "mother city". A directly related Greek word is *D¶m¶t¶r*, which names the goddess of the harvest. Other words that derive from this root, via the Latin word *mâter*, include ‘maternal’, ‘matron’, ‘matrimony’, ‘matrix’, ‘matter’, and of course ‘mother’.
[11] There are many more highly specialized examples as well.

thermo*meter*, baro*meter*, anemo*meter*	weather measurement devices
speedo*meter*, odo*meter*, tacho*meter*	on most auto instrument panels
parking meter, taxi meter, electric meter	various measuring devices
voltmeter, ohmmeter, etc.	various measuring devices
dia*meter*, peri*meter*	characteristics of a circle
para*meter*	"beside measure"
millimeter, centimeter, meter, kilometer	units of length/distance
pentameter, hexameter, etc.	measure how many "feet" in a line of poetry
meter	as in music: 4/4, 3/4, 6/8, etc.
geo*metry*	measures the earth
trigono*metry*	measures tri-gons (i.e., triangles)
opto*metry*	measures optic stuff (eyes)
sym*metry*	"same measure"

3. Definitions of Measurement

For the sake of comparison, we consider several definitions of ‘measurement’ gleaned from various sources, including ones devoted to science and philosophy In each case, we underline words to call attention to a noteworthy feature of the definition.

1. American Heritage Dictionary[12]

   [To measure is] to ascertain the dimensions, quantity, or capacity of..

2. Dictionary of Consciousness[13]

   A measurement is an objective procedure the purpose of which is to determine a number that is characteristic of a specified physical situation.

3. Free On-Line Dictionary of Computing[14]

   [To measure is]to ascertain or appraise by comparing to a standard; to apply a metric.

4. Web Dictionary of Cybernetics and Systems[15]

   [Measurement is] the process of ascertaining the attributes, dimensions, extent, quantity, degree or capacity of some object of observation and representing these in the qualitative or quantitative terms of a data language.

5. R. McCleary (Web page at UCal-Irvine for course on Research Design)[16]

   Measurement is the process of assigning each of *n* individuals or elements to one and only one of *k* categories.

[12] http://www.bartleby.com/61
[13] http://www.ec3.com/Upperized/dictiona.htm
[14] http://foldoc.doc.ic.ac.uk/foldoc/index.html
[15] http://pespmc1.vub.ac.be/ASC/indexASC.html
[16] http://mrrc.bio.uci.edu/se10/measurement.html

4. An Initial Account of Measurement

According to an often cited account of S.S. Stevens[17], a special committee of the British Association for the Advancement of Science spent eight years (1932-40) discussing the possibility of measuring sensory events, but they never reached a consensus, so divided were they on what it means to make a measurement.

Although it is doubtful that we could produce an account that would satisfy every member of this special committee of BAAS, we nevertheless wish to propose an initial general account of measurement, to serve as a starting point in our discussion.

> initial account of measurement (ccc):
>
> a measurement is a procedure for the purpose of
> ***c****omparison*, ***c****lassification*, and ***c****ommunication*

Now, apart from the alliterative quality of this account, which aids memory, we need to assess whether it even comes remotely close to accounting for what measurement is.

5. What is a Comparison?

We begin by considering the connection between measurement and comparison. Let's first make sure we have a reasonably clear and precise idea what a comparison is. We take the following dictionary definition (from *AHD*) as our starting point.

> **2.** To examine in order to note the similarities or differences of.

Etymology might be helpful here. First, the word 'compare' combines two Latin morphemes – *com* [with] and p^3r [equal]. The latter word also directly gives rise to the word 'par', as well as the related words 'part', 'parse', and 'parcel'. The word 'par' in turn has a variety of uses pertaining to comparison against an established standard. For example, probably the most popular use of 'par' pertains to golf, in which every hole on a given golf course has a par value, which is presumably a *standard* against which golfers compare/measure themselves.[18]

Next, we note that the above definition is not entirely satisfactory logically speaking, since it is not obvious from the definition what the exact grammatical category of 'compare' is. Let us consider this briefly. First, it seems that 'compare' is a transitive verb, since 'examine' is a transitive verb. However, whereas the following is grammatical

> the biologist examined the swan

the following does not seem to be grammatically admissible.

> (*)[19] the biologist compared the swan

---

[17] S.S. Stevens, "On the Theory of Scales of Measurement", *Science*, Volume 103 (1946).

[18] Ordinarily, one wants to be at par or above par. However, in golf, one wants to be at or below par, since one is trying to sink the ball in the hole using the fewest strokes.

[19] It is customary for linguists to place an asterisk next to phrases they count as ungrammatical.

In other words, ‘compare’ is not a *simple* transitive verb like ‘examine’ which takes a *single* grammatical object. Rather, it is more like ‘recommend’, ‘buy’, and ‘sell’, which take both direct objects and indirect objects. Thus, the following seem grammatically plausible, if not scientifically plausible.

Jay compared the swan *with* his watch
Kay compared the swan *to* her watch

This brings up an interesting usage issue – what is the difference between ‘to’ and ‘with’? According to *AHD*, ‘compare’ usually takes the preposition ‘to’ when the comparison involves ostensibly *unlike* things; for example, a poet might compare his true love **to** a summer day. However, not all metaphor is poetic. Scientists sometimes compare the human brain **to** a computer, and they sometimes compare chemical atoms to the solar system.[20] Next, according to *AHD*, ‘compare’ usually takes the preposition ‘with’ when the comparison involves ostensibly *like* things. For example, the police may compare a fingerprint found at a crime scene **with** fingerprints on file in their database. Notice, in this connection, that it is also acceptable to use the preposition ‘against’.

In this connection, we observe that ‘compare’ also has a usage as an intransitive verb as in the following example.

modern novels simply do not *compare with* the classics[21]

This usage of ‘compare’ leads to the following odd expression.

modern novels *cannot* be compared with classics

This is literally false in at least one sense. We *can* compare modern novels to the classics, and when we do we find that the modern novels are not nearly as good! Similarly, if we say something is incomparable, we are saying that nothing compares *favorably* with it, which means that it is significantly better (or bigger, or more important) than anything else in the relevant category.

The considerations so far suggest that the relevant usage of ‘compare’ treats it as a transitive verb with two objects, neither of which is optional.[22] The basic form is:

p compares x and y

where x and y are the objects being compared with each other, and p is presumably an agent or the instrument of an agent.[23] Alternative forms are:

p compares x with y
p compares x against y

However, this can’t be the whole story. When we compare two objects, we don’t compare them *simpliciter*; rather, we compare them with reference to some *standard* or *mode* of comparison (what is sometimes called a "dimension"). If I ask you how Jay compares with Kay, you would naturally ask me

[20] Both analogies have their limitations.
[21] Note that the following expresses a similar sentiment.
modern novels simply do not *measure up to* the classics
[22] By contrast, the verbs ‘buy’ and ‘sell’ have optional indirect objects. I can sell my car (to Smith); I can buy my car (from Smith).
[23] In philosophy, an agent is something that acts, where an act is usually regarded as voluntary. Opening a door is an act; being hit by the door as it opens is not an act.

in return “with respect to what standard?” We can compare two persons with respect to height, weight, age, wealth, intelligence, etc. Thus, it appears that the *underlying* concept is more properly understood as a four-place relation, expressed by sentences of the form

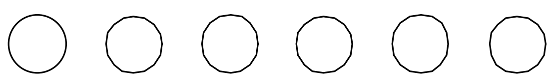
p compares x and y *with respect to* d

where d is the mode of comparison, or dimension.

6. Is Every Measurement a Comparison?

Our initial hypothesis is that measurements involve comparison, classification, and communication. Let us now consider the first component. The question is: does every measurement involve a comparison? We need to examine measurements and check whether they all involve comparisons.

At this point, we have discussed only one clearly marked example of measurement – measuring the size of a set. So the natural question is – when we measure the size of a set, are we making a comparison? For example, when we count how many people are in the room, are we making a comparison? If so, to what are we comparing the people in the room? In this connection, recall that fundamental to counting is the notion of *one-to-one correspondence*.[24] In particular, two sets are said to be *equally-big* if there exists a one-to-one correspondence between them. Also recall that when we count a set S, we establish a one-to-one correspondence between S and an appropriate set of initial numbers.[25] For example, in counting the following marbles, we establish a one-to-one correspondence between the marbles and the numbers one, two, three, four, five, and six.



one two three four five six

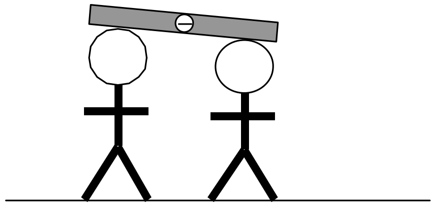
Thus, counting is fundamentally comparison of sets. This datum accordingly provides positive evidence, although *not conclusive* positive evidence, for our hypothesis that every measurement is a comparison.[26]

7. Is Every Comparison a Measurement?

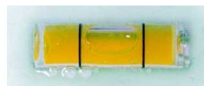
Now, clearly one datum is not conclusive; we need to examine many other examples of measurements. But before we do that, let us briefly examine a related question – is every comparison a measurement?

What are some examples of comparison? A very simple example of a physical comparison is when we place two people next to each other – say back to back with their heads nearly touching – in order to ascertain which one is taller. This can be made more precise by using a carpenter’s level, in which case we can compare the two people’s heights using the technique graphically depicted as follows.

[24] See “Numbers and Counting”.
[25] Or numerals! For these purposes, it does not matter whether we use numbers or numerals; that is the beauty of counting!
[26] See chapter on confirmation.



In this picture, the stick figures represent the two people being compared, and the rectangular shape represents the measuring instrument. The latter may be presumed to be a carpenter's level, also called a "spirit" level.[27] Such instruments have one or more tubes in them, one of which is pictured as follows.



The tube contains alcohol, or ether, or a similar fluid, in which there is a bubble. The idea is that the bubble shifts to the higher side of the level, thereby indicating who is taller (presuming they are standing on a level surface). In the picture above, it is apparent that the person on the left is taller than the person on the right.

Is this an example of a measurement? Unlike counting, this procedure does not produce a specific "number" as an output. Rather, it only produces an assessment or judgment concerning who is taller than whom. So, perhaps it does not count as a *paradigmatic* example of measurement. Nevertheless, it does not seem unreasonable or farfetched to say that the described procedure constitutes a measurement in which we measure each person *against* the other.

Let us compare this example with our previous example involving set sizes. Underlying the concept of counting is the comparative concept "bigger". We learn to compare set-sizes before we learn to do this using an intermediate standard (the counting numbers). For example, we can recognize that we have equally-many fingers on each hand without recognizing exactly how many fingers this is. Similarly, we can recognize that one set is bigger than another (e.g., there are more people than fish), without recognizing the exact quantitative difference. These comparisons are aided by introducing intermediate standards (the numbers), which provide a technique for comparing sets.

In light of these considerations, we introduce the following two distinctions.

(1a) A measurement of an object is said to be *standard* precisely when it compares the object against a "standard" object.

(1b) A measurement of an object is said to be *non-standard* precisely when it is *not* standard

(2a) A comparison of two objects is said to be *indirect* precisely when they are compared by way of an intermediate (third) object.

(2b) A comparison of two objects is said to be *direct* precisely when it is not direct.

What qualifies as a standard object is largely a matter of social/linguistic convention, where the social/linguistic unit is also partly a matter of context. This is because the chief purpose of standard measurement is communication, and communication involves many contextual elements. Let us illustrate this by considering a simple example. Suppose I tell you that my dog is small. What do I mean? How much information have I conveyed to you? This depends upon the context within which

[27] Spirit levels were invented in France in the 17th Century.

the assertion is made, which includes the social/linguistic factors pertaining to the two of us at the time of the utterance. For example, what are we discussing size in relation to? The particular breed[28]? dogs? pets? domesticated animals? mammals? things in general? Even granting that the context is dog-sizes, whether communication takes place depends upon whether (and to what extent) you and I share a common standard for what counts as a small dog. If I tell you my dog weighs 17 pounds, that will provide you with information, *provided* you know how much a *typical* dog weighs *in pounds*. Alternatively, we have to come up with some common types of dogs (or things) with which we are both familiar, and I have to tell you how my dog compares to these types. For example, I could say that my dog is smaller than a cocker spaniel, but bigger than a toy poodle, supposing that you have passing knowledge of these two other breeds.

The distinction between *direct* and *indirect* comparison is closely related. A direct comparison is illustrated by placing two objects next to each other to compare their heights. By contrast, an indirect comparison occurs when we compare two objects by way of an intermediate object, for example, when we measure each object against a standard. For example, when we compare two sets directly against each other, we are engaged in direct comparison. On the other hand, when we count the members of each set, and we compare the resulting numbers, we are engaged in indirect comparison.

8. Assessment and Measurement

In the previous section, we considered comparing the heights of persons p_1 and p_2. The idea is that a direct comparison of p_1 and p_2 results in a judgment of which of the following is the case.

(1) p_1 is taller than p_2
(2) p_2 is taller than p_1
(3) p_1 and p_2 are equally tall

This seems to be a special case of judgment or assessment.

By way of comparison, let us consider another example of assessment. Suppose you are camping in the mountains, and one morning you wake up only to discover a bear staring at you. At this point, the two of you are going to "size each other up", where "sizing up" is presumably a rudimentary form of measurement. Now obviously you don't take a tape measure to the bear. Rather, you *assess* the bear in relation to you, the situation, and your beliefs about bears. Also obviously, you don't assess the bear as a potential philosophical interlocutor or chess opponent! No, you probably assess the bear as a potential threat to your physical well-being. In this connection you consider several questions: Is it a grizzly bear or (merely!) a black bear? Is it a nursing mother with cubs near by? Is it behaving aggressively? Do you feel macho? What should you do? Do you shout at the bear, trying to scare it away? Do you run for the nearest tree? Do you curl up and play dead? Each of these actions is worth considering, although you probably should do your calculations fairly quickly.

Assessment is a kind of measurement. Is every assessment a comparison? What does it mean to assess? What sort of thing can I assess? The logically simplest sort of assessment, which serves as a model for all examples, is the following.

assessing whether (or not) Φ

[28] The breed in question is Cavalier King Charles Spaniel, a dog breed named after King Charles II, who was supported by the royalist group known as the Cavaliers. The royalists were eventually able to overthrow Oliver Cromwell, who had earlier overthrown Charles I, and restore Charles II to the throne. In any case, King Charles II was very fond of this particular breed of small spaniels, and had dozens of them.

where Φ expresses a simple proposition. A simple proposition involves a single predication, which in turn concerns one or more individuals.[29] For example, I can assess whether a given swan is white. Is this a comparison? Well, a comparison is between one object and another. When we assess whether a swan is white, what are we comparing? Over the years, we gradually learn how to apply the adjective 'white'. Each time we successfully apply it, we augment our continually growing mentally internalized standard of what sorts of objects the term 'white' refers to. So, when we come upon a novel object, in order to assess whether it is white, we *compare* that object against this mental standard. Of course, it would be nice to carry an actual physical standard of whiteness (for example, freshly fallen snow!) But this is impractical, so we have to rely on our memory.

9. Measurement as Classification

In the previous chapter, we discussed measurement as assessment. In the present chapter, we pursue this idea more generally. The general form of assessment is given as follows.

assessing whether Φ_1, …, or Φ_k

where Φ_1, …, Φ_k constitute k-many choices, where it is presumed that each choice excludes all the others. In other words, the 'or' is an exclusive 'or'. The simplest example is when there are two choices; an example is judging whether a swan is white or not. The next simplest example is when there are three choices; an example is comparing the heights of two people.

Based on this idea, we can provide a definition of measurement as follows.

a measurement is a *procedure* by which one *assigns* a given *individual* in a *domain* to exactly one *category* in a collection of categories $\{K_1, \ldots, K_m\}$.

Example 1: when examining a bird, assessing whether the bird is, or is not, black; in this case, the domain consists of birds, and the categories are $\{B, \sim B\}$.

Example 2: when comparing a person against a standard object, assessing whether the person is taller, shorter, or equal in height to the standard object; in this case, the domain consists of persons, and the categories are $\{T, S, E\}$.

Example 3: when comparing a person against a standard object, assessing how much bigger/smaller the person is than the standard; in this case, the domain consists of persons, and each category corresponds to a positive fraction (1/2, 2/1, 2/3, 3/2, etc.)[30]

[29] In most cases. Sometimes, there is only a dummy subject, as in 'it is raining'. This is probably best understood as a zero-place predication, since the term 'it' does not refer to anything.
[30] See chapter "Other Numbers" for a discussion of fractions.

10. Scales of Measurement

It is customary to refer to the categories $\{K_1, \ldots, K_m\}$ as a *measurement scale*. The word ‘scale’ has three distinct entries in the dictionary, corresponding to three different origins. One entry derives from the French word ‘*escale*’ [shell], and pertains to fish and reptiles whose skin if often covered with scales. Another entry derives from an old Norse word ‘*sk*^{3}*l*’ [bowl, balance], and pertains to instruments used to weigh things (e.g., fish). The third entry derives from the Latin word ‘*sc*^{3}*lae*’ [ladder], and pertains to a system of graduated categories, perhaps the simplest example being a musical scale (e.g., do, re, mi, fa, so, la, ti, do). The notion of measurement scale derives from the latter use of the word ‘scale’.

At the moment, we are primarily interested in simple (*one-dimensional*) measurement scales. Following Stevens[31], simple measurement scales are usually classified into four basic types.

(1) **nominal** scales
(2) **ordinal** scales
(3) **interval** scales
(4) **ratio** scales

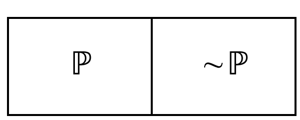
11. Nominal Scales

Every measurement scale consists of a collection $\{K_1, \ldots, K_m\}$ of categories, into which the objects in the relevant domain are classified. A nominal scale is the simplest possible type – on in which the categories have no non-trivial topological or metrical relations to one another. In particular, even if the categories are labeled by numbers, these numbers carry *no quantitative information* except to distinguish one category from another.

For any given domain, there are two trivial nominal scales associated with that domain – the universal scale, and the identity scale. Whereas the *universal scale* classifies every object into the same category, the *identity scale* classifies each object into its own unique category. Such scales are largely useless, and are included merely for mathematical completeness.

The simplest non-trivial nominal scale classifies objects into two categories; these scales are variously called *bi-nominal*, *bi-partite*, and *two-valued*. For example, every one-place predicate $\mathbb{P}$ gives rise to an associated bi-nominal scale, which classifies objects according to whether they are $\mathbb{P}$ or not $\mathbb{P}$. Another simple example of a nominal scale is the classification of humans into male and female. Along similar lines, some languages – including French, Spanish, and Italian – classify nouns into two genders – masculine and feminine. Other languages – including Latin and German – classify nouns into three genders – masculine, feminine, and neuter. A nominal scale with three categories is variously called *tri-nominal*, *tri-partite*, and *three-valued*. Another example of a tri-nominal scale is the classification of U.S. voters into democrats, republicans, and independents.

Each nominal scale can be graphically depicted as a box divided into non-overlapping regions. The following are examples.



[31] S.S. Stevens, “On the Theory of Scales of Measurement”, *Science*, Volume 103 (1946).

male	female

masculine	feminine	neuter

democrat	republican	independent

Oftentimes, a classification system has a "miscellaneous" or "none of the above" or "other" category, in which we place all those items that don't fit into one of the major categories. For example, we might classify humans according to their religious affiliation, in which case we might employ the following five major categories, plus an "other" category.

Buddhist	Christian	Hindu	Jewish	Muslim	Other

The key fact about nominal scales is that their categories have *no topological features* – which is to say that there is no up-down, or right-left, or bigger-smaller, or even "in between"; there is only "difference".[32] For example, in the above classification of religious affiliations, the placement of Jewish between Hindu and Muslim has no objective significance. The pictorial arrangement of the categories is arbitrary; we happened to have placed the categories in alphabetical order. Any other arrangement would be equally valid. When we classify objects using a nominal scale, all we can presume about these categories is that objects in one category are *different* from objects in every other category. In particular, we cannot presume that this difference can be further delineated.

12. Ordinal Scales

If the categories of a measurement scale are organized so that it makes sense to say that one category is smaller than another category [or to the left of, or below, or before, …], then we have what is called an *ordinal scale*. More abstractly speaking, an ordinal scale is a classification in which the categories are linearly-ordered.[33] Recall from an earlier chapter that a *linear-ordering* may be defined as a relation – generically depicted by '$>$' [or '$<$'] – satisfying the following conditions.

(c1) $x > y \;\rightarrow\; y \not> x$ [asymmetry]
(c2) $x > y \;\&\; y > z \;.\!\rightarrow\; x > z$ [transitivity]
(c3) $x \neq y \;\rightarrow\!.\; x > y \vee y > x$ [connectivity]

Here, the shorthand '$\not>$' is defined in the usual way.

(d1) $x \not> y \quad =_{\text{df}} \quad \sim[x > y]$

It is customary to define a converse relation $<$ in the obvious manner, as follows.

[32] In the mathematical discipline of Topology, such a space is called a "discrete" topological space. These don't count as topologically interesting.

[33] Once again, we are concentrating on one-dimensional measurement scales. For multi-dimensional measurement, linearity does not make sense, and must be replaced by a concept of multi-linearity.

(d2) $x < y \quad =_{\text{df}} \quad y > x$

The resulting relation is easily seen to be a linear-ordering as well.

Probably the best-known example of an ordinal scale is given by the letter grades – A-F. What makes this classification an ordinal scale rather than simply a nominal scale is that the categories are regarded as ordered. Getting an A is not merely *different* from getting a B, or C, …; it is *better*! This can be depicted as follows.

A > B > C > D > F

Some grading systems employ more categories; for example, at UMass, there are three extra intermediate categories, which produces the following scale.

A > AB > B > BC > C > CD > D > F

Other schools use an even more precise classification, using *pluses* and *minuses*. For example, Mount Holyoke College and Pomona College use the following grading scale.

A > A– > B+ > B > B– > C+ > C > C– > D+ > D > D– > F

We can also go the other direction, and produce a less precise scale, one in which the only categories are "pass" and "fail", as follows.

P > F

Another example of an ordinal scale comes from the natural world, where scientists have discovered a species of slime mold[34], *Physarum polycephalum*, which has evolved a system of sexual reproduction that is truly bizarre. In particular, this species has thirteen distinct sexual castes[35], rather than just the usual two. I use the word 'caste' intentionally, since the categories are ordered in terms of power; in particular, any individual of a given caste can donate genetic material to any individual of a lower caste, but not vice versa. The following diagram depicts the ordering using Greek letters.[36]

$\alpha > \beta > \gamma > \delta > \varepsilon > \zeta > \eta > \theta > \iota > \kappa > \lambda > \mu > \nu$

Note carefully that this order is not a matter of artificial social organization, which would be absurd to posit in the realm of slime molds; rather, it is purely biological.

There are of course genuine examples of social hierarchies. For example, wolf packs are usually organized so that there is an alpha-male, an alpha-female, a beta-male, a beta-female, and after that the rest of the pack may simply be labeled as gammas. This is probably the inspiration for the terminology in Aldous Huxley's *Brave New World*, which portrays a fictional human society that is organized into

[34] Slime molds are currently thought to be one of the most primitive (i.e., earliest) organisms in the eukaryote realm (super-kingdom).

[35] Alun Anderson, "The Evolution of Sexes", *Science*, New Series, Vol. 257, No. 5068. (Jul. 17, 1992), pp. 324-326.

[36] The letters are respectively – alpha, beta, gamma, delta, epsilon, zeta, eta, theta, iota, kappa, lambda, mu, nu, …

three castes – alphas, betas, and gammas. Another example of a linear social order occurs in barnyard chickens, who organize themselves into a pecking order; in particular the alpha-chicken gets to eat (i.e., peck) before the beta, who gets to eat before the gamma, etc.

13. Quick Review

For the moment, let us concentrate on tri-partite classifications, which we label as alphas, betas, and gammas. In a nominal scale, these categories represent qualitative differences between individuals in the domain. This can be represented by the following graph.

$$\boxed{\alpha}$$

$$\neq \qquad \neq$$

$$\boxed{\beta} \neq \boxed{\gamma}$$

This means that alphas are different from betas and gammas, and betas are different from gammas, but there is no presumed inherent order to these categories.

On the other hand, in an ordinal scale the categories reflect an inherent ordering of the individuals, as depicted in the following graph.

$$\boxed{\alpha} > \boxed{\beta} > \boxed{\gamma}$$

In other words, alpha is not merely *different* from beta; alpha is (in some sense) *better* than beta. Similarly, beta is not merely different from gamma; beta is *better* than gamma.

14. Comparing Differences

In an ordinal scale, we can meaningfully rank the categories. For example, in a society with three castes – α, β, γ – it is better to be an alpha than a beta, and it is better to be a beta than a gamma. The obvious question is: how much better? For example, suppose you are promoted from gamma to beta, and later you are promoted from beta to alpha. Which promotion is the "bigger" promotion? Alternatively stated, which constitutes a bigger improvement?

By way of answering this question, let us introduce differential notation as follows.

$$\alpha\text{–}\beta \qquad \alpha\text{–}\gamma \qquad \beta\text{–}\gamma$$

The question then is: can we compare caste-differences? Logic alone partly answers the question, by providing the following results.

$$\alpha\text{–}\gamma > \alpha\text{–}\beta$$
$$\alpha\text{–}\gamma > \beta\text{–}\gamma$$

The above comparisons seem to be logically inherent to the concepts of "better" and "improvement". For example, the following seems clearly to be a valid argument.

being an *A* is better than being a *B*
being a *B* is better than being a *C*
therefore, changing from *C* to *A* constitutes a bigger improvement than changing from *C* to *B*, and also a bigger improvement than changing from *B* to *A*.

On the other hand, it is not obvious *a priori* what the answer is to the following question.

how do α–β and β–γ compare?; which difference (if either) is bigger?

The possible answers include the following.

(a1) it makes **no sense** to compare differences;
(a2) it makes sense to compare differences, but **only qualitatively**;
(a3) it makes sense to compare differences **quantitatively**;

If we posit answer (a1), we in effect maintain that the scale is "incorrigibly" ordinal; it cannot be further elaborated. On the other hand, if we posit (a2) or (a3), we in effect maintain that the scale can be further elaborated.

15. Qualitative Difference-Comparison

If we posit answer (a2), then we can compare α–β and β–γ qualitatively, which simply means that the question "how do α–β and β–γ compare?" is answered by one of the following.

α–β > β–γ
α–β < β–γ
α–β ≡ β–γ

We might wish to graphically depict these judgments respectively as follows.

$$\boxed{\alpha} \;\gg\; \boxed{\beta} \;>\; \boxed{\gamma}$$

$$\boxed{\alpha} \;>\; \boxed{\beta} \;\gg\; \boxed{\gamma}$$

$$\boxed{\alpha} \;>\; \boxed{\beta} \;>\; \boxed{\gamma}$$

Here, the symbol '>>' and its derivatives ('>>>', etc.) are used to compare improvements *qualitatively*. For example, $\alpha \gg \beta$ indicates a bigger difference than $\alpha > \beta$. Note carefully that we do not want to impute any further significance to the symbol; for example, we don't mean that the β-α difference is *twice* as big as the γ-β difference.

16. Quantitative Difference-Comparison – Interval Scales

If we wish to say that one difference is *twice* as big as another difference, then are discussing quantitative comparison of differences, which brings us to the notion of an *interval scale*.

For example, consider the usual letter grades – A, B, C, D, F. How do grade-differences compare? By logic alone (see above), the following seem evident.

A–F	>	A–D	>	A–C	>	A–B
				A–C	>	B–C
		A–D	>	B–D	>	C–D
				B–D	>	B–C
A–F	>	B–F	>	C–F	>	D–F
				C–F	>	C–D
		B–F	>	B–D	>	B–C
				B–D	>	C–D

For example, as a simple matter of logic, the difference between A and F is bigger than the difference between A and B. On the other hand, it is not obvious *a priori* what the answer is to the following question.

how do A–B and B–C compare?; which difference (if either) is bigger?

There is a standard answer to this question. For example, academic administrators have traditionally postulated *ex cathedra* that the difference between any two adjacent letter grades is equal to the difference between any other two adjacent letter grades. Thus, by hypothesis:

A–B ≡ B–C ≡ C–D ≡ D–F

Moreover, it is postulated *ex cathedra* that differences can be meaningfully added so that:

A–C ≡ A–B + B–C
A–D ≡ A–C + C–D
etc.

These postulates are not made explicitly and openly, but rather implicitly and surreptitiously. These postulates are smuggled into our conceptual framework by the simple and seemingly harmless practice of assigning a grade-point to each letter grade – 4 to A, 3 to B, etc. This is conceptually harmless in itself, so long as we treat the numbers 4, 3, 2, 1, 0 as merely an ordered set of labels/categories (i.e., an ordinal scale). If they are merely labels (like addresses), then they cannot be meaningfully added or multiplied. But, it is precisely here that the mischief occurs. Once the letter grades (A,B,C,D,F) are converted to numbers (4,3,2,1,0), one is tempted to construct further quantitative measures. Thus arises the statistical chimera known as the grade-point average (GPA). This is not a mere figure, since decisions regarding a student's academic standing are made based on this "measure". For example, at many universities and colleges a student is required to maintain a GPA of at least 2.0.

We are not suggesting that the GPA is completely without merit. Rather, we are simply pointing out that its use is based on assumptions that need to be made explicit. In particular, taking averages of grades presupposes that differences between the letter grades are quantifiable in a certain way, listed above.

Matters get even more out of hand when we consider augmentations of the basic five letter grades. For example, the eight-grade system usually presumes that the intermediate letter grades – AB, BC, CD – are exactly intermediate between the surrounding grades. Accordingly, they are assigned point-values of 3.5, 2.5, and 1.5, respectively, which may be graphically depicted as follows.

A	>	AB	>	B	>	BC	>	C	>	CD	>	D	>>	F
4		3.5		3		2.5		2		1.5		1		0

Here, a single '>' represents one step, and '>>' represents two steps. In other words, the underlying presupposition is that the difference between D and F is twice as big as the difference between any other two adjacent grades.

Another numerical scheme is obtained if we insist that all the steps are equal, in which case the simplest numerical scale would be as follows.

A	>	AB	>	B	>	BC	>	C	>	CD	>	D	>	F
7		6		5		4		3		2		1		0

Encountering a seven-point grading scale for the first time is a bit unsettling. It is very similar to encountering a metric measurement (for example, in centimeters); one's immediate instinct is to convert the measurement into a known quantity (for example, inches).[37] And indeed, one is naturally inclined to apply a scaling factor to a seven point scale – multiply by 4/7 – to bring it in line with one's familiar scale. But notice that this equal-division scale does not produce the same GPA as the standard scale, which can be seen by comparing the following to the standard table above.

A	AB	B	BC	C	CD	D	F
4	3.43	2.86	2.28	1.71	1.14	.57	0
4	3.5	3	2.5	2	1.5	1	0

Which of these scales is correct? It depends upon the answer to the following questions.

(1) is the difference between D and F *twice as big* as the difference between A and AB?
(2) is the difference between D and F *equal to* the difference between A and AB?
(3) neither of the above!

I submit that the answer is (3). Rather, the assignment of numbers to letter-grades is largely arbitrary. Nevertheless, it is a tradition that would be difficult to overthrow. It has legal standing, but not mathematical standing.[38]

Whereas a grade-point average may not be a legitimate statistical measure for letter grades, the median-grade is considered by statisticians as a legitimate measure. The median grade is the grade that represents the half-way point. For example, if your grades are given as follows, ordered top to bottom.

A A AB B B BC BC C CD D

then the median grade is B, since half of the grades are B or better.

---

[37] I am speaking from the viewpoint of most Americans. The opposite would be true for just about everyone else in the world.

[38] A similar situation occurs in voting in the U.S. The U.S. has a legally binding procedure by which a candidate wins an office. This is far better than other methods of political change that the world has experienced. Nevertheless, the political question is completely separate from the question whether our particular voting procedure is logically sound.

17. Standard Examples of Interval Scales

It is not completely obvious that letter grades form an interval scale. What is a legitimate example of an interval scale? First, it is generally accepted that time ("when") can be measured by an interval scale, where the intervals correspond to temporal durations ("how long"). In particular, temporal durations can be meaningfully added. The two most famous historical examples of interval scales are the everyday temperature scales – the Celsius scale and the Fahrenheit scale. We will investigate both time measurement and temperature measurement in later sections.

18. Ratio Scales

In an interval scale, it makes sense to add differences. For example, although it is dubious, let us "get with the program" and accept letter grades as a genuine interval scale. Then, we can say the following.

$$\begin{array}{ccccc} \text{A–C} & \equiv & \text{A–B} & \equiv & \text{B–C} \\ \text{A–C} & \equiv & \text{A–B} & + & \text{B–C} \end{array}$$

So, we can say that the difference between an A and a C is *twice* the difference between A and B.

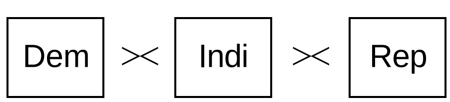
A ratio scale is based on this exact idea, except that the ratios[39] apply, not just to category-differences, but also to the categories themselves. In particular, it makes sense to say that one category is twice (or half) as big as another category.

All the best scales are ratio scales, which include the measurement scales for length and weight. The key feature of a ratio scale is that it has a "natural zero"; accordingly, a ratio scale does not have negative values. For example, it makes sense to talk about zero length and zero weight, but it does not make sense to talk about negative length or negative weight.

Note that every interval scale gives rise to an associated ratio scale. For example, although time ("when") does not constitute a ratio scale, duration ("for how long") does. More about time measurement later.

19. Quasi-Ordinal and Quasi-Interval Scales

In addition to the official scales promoted by Stevens, there are some other types of scales worth considering. In particular, the ordinal and interval scales give rise to associated quasi-scales – the quasi-ordinal scale, and the quasi-interval scale. The basic idea is that a quasi-scale is a scale minus its *orientation*. Although there is no inherent left-right or up-down, there is nevertheless an inherent "between" relation. The simplest example is ⟨Democrat, Independent, Republican⟩. Although there is no inherent order to these categories, there is an important sense in which Independent is *between* Democrat and Republican.[40] This can be represented graphically as follows.



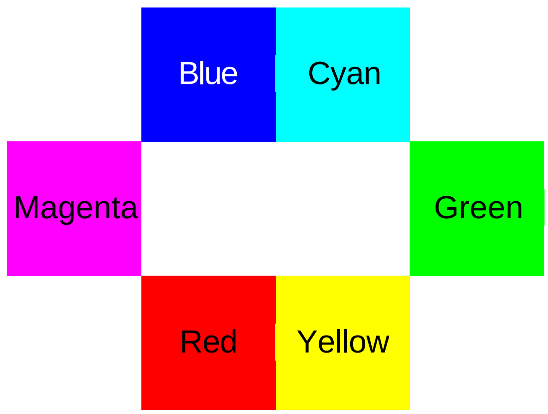
[39] Recall material on fractions in "Other Numbers".
[40] Someone may object that there is a left-right orientation to political parties, but I would say that this terminology is largely inaccurate and simplistic, that it says more about our preoccupation with simple spectral classification than it says about political reality.

Another example consists of the so-called spectral colors – red, orange, yellow, green, blue, indigo, violet, which can be graphically represented as follows.[41]

R >< O >< Y >< G >< B >< I >< V

Once again, saying that red is bigger than violet or saying that violet is bigger than red makes little or no sense.[42] On the other hand, there is an important sense in which orange is between red and yellow, yellow is between orange and green, etc.

A more dramatic example of a non-ordinal scale is obtained by considering the standard color wheel used to teach simple color theory to artists and graphic designers. According to the most fundamental classification of colors, there is a basic circle of six colors identified as follows.



We can also depict it as follows.

M >< B >< C >< G >< Y >< R >< M

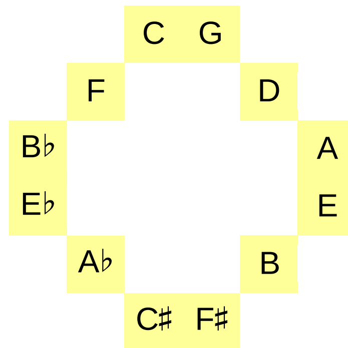
Note that magenta appears twice in this representation in order to indicate that the color spectrum folds back on itself.

Whereas a quasi-ordinal scale is an ordinal scale minus orientation, a quasi-interval scale is an interval-scale minus orientation. For example, location on a line is best understood as a quasi-interval scale, since it makes little sense to say that one location is bigger, or smaller, than another location.

Another example of a quasi-interval scale is the "circle of fifths" in music which is an arrangement of scale degrees according to musical adjacency (see chapter on musical scales). The simple version of the circle of fifths is given as follows.

[41] This catalog of colors is a bit artificial, even a bit superstitious. That there are seven such colors, rather than six, is based primarily on the early pre-occupation with the number seven (even by intellectual giants like Newton!). Most people who have been to American elementary schools can recite the famous "Roy G. BIV", but very few people (if any!) can say authoritatively what color indigo is, and very few people (if any!) can identity the indigo portion of a color spectrum.

[42] Someone may object that violet photons are individually more energetic than red photons. My response is that I am not talking about photons, but about colors, which are not photons. To be sure, photons of various wavelengths are involved in the production of color sensations, but they are not themselves colors. Similarly, knives and guns are occasionally involved in the production of pain sensations, but they are not themselves pains.



The basic idea is that each scale degree ("key") has two closely related scale degrees, which are adjacent. For example, C-major is between F-major and G-major, G-major is between C-major and D-major.

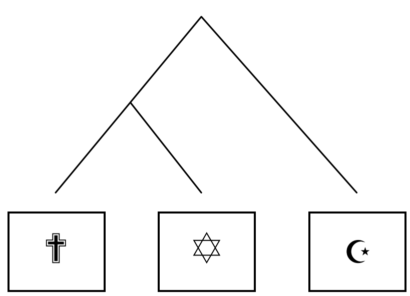
20. Absolute Scales

Another scale-type not propounded by Stevens, but which is worth considering is the scale at the extreme top – absolute scales. An absolute scale is like a ratio scale in which only whole numbers are allowed. Alternatively, an absolute scale is one in which there is only one admissible "unit", with respect to which all measurements yield whole numbers. The comparison of set sizes provides the only obvious example of such a scale. For example, when we say that there are three swans in the pond, this is relative to an understood and absolute unit.

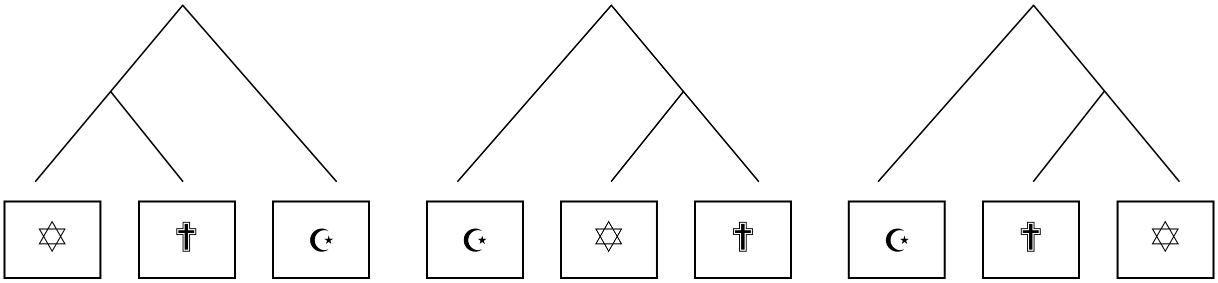
21. Differential-Comparison Revisited

In an earlier section, we examined whether one can compare *ordinal* differences, and based on this we arrived at the notion of an interval scale. We now back up a bit and consider whether we can compare simple differences – for example, the differences that appear in a nominal scale. In this connection, we consider the question whether it might make sense to say that two categories resemble each other more than they resemble a third category; alternatively, we can say that the third category is more different than the first two categories.

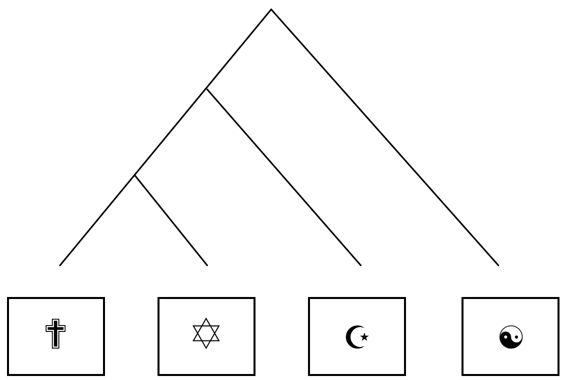
For example, consider religious affiliation. Although there does not seem to an inherent order to these categories, it is generally agreed by comparative religion scholars that Christianity and Judaism to are more similar to each other than either is to Islam. This differential-difference gives rise to a hierarchical organization that looks something like the following.



Note carefully that the drawing suggests that ✡ is adjacent to ☪, but ✝ is not. This is an unfortunate artifact of the representation. In fact the arrangement is that ✡ is just as far from ☪ as ✝ is. It may help to think of this sort of graph as a picture of a mobile (one of those dangling light weight sculptures), in which the categories hang freely below various pivot points. So, both ✝ and ✡ can pivot freely around each other, and the ✝-✡ pair can pivot freely around ☪. So the following diagrams represent the situation equally well.



Next, we observe that we can enlarge the above classification to include other major religions. For example, most scholars of comparative religion would agree that these three religions resemble each other more than any one of them resembles Buddhism, in which case we have a larger classification system, depicted as follows.[43]



Bear in mind that the various terminal nodes are free to pivot about common focal points. So although it appears that ☪ is adjacent to ☯, it is in fact just as far from ☯ as ✝ and ✡ are.

This is clarified a bit when we examine the logic of comparative-similarity, which is quite complex, since it is a three-place relation. Let us combine the symbols '$\circ$' and '$<$' into a symbol-complex in such a way that:

$a \circ b < c$ =: a and b resemble each other more than either resembles c[44]

What principles govern this predicate? The following are postulated.

---

[43] I do not have an easily recognized icon for Buddhism, so I have simply used a yin-yang symbol.
[44] The latter expression can be stated more succinctly as follows:

> b is more like a than c

provided we take advantage of the usual ambiguity of 'more than', so that this means *both* of the following.

> b is more like a than b is like c
> b is more like a than c is like a

For example, the sentence

> Jay likes Kay more than Elle

is ambiguous between

> Jay likes Kay more than Jay likes Elle

and

> Jay likes Kay more than Elle likes Kay

(1)	$x \circ y < z \;\rightarrow\; y \circ x < z$	[symmetry of $\circ$]
(2)	$a \circ b < c \;\rightarrow\; a \circ c \not< b$	[asymmetry of $<$]
(3)	$a \circ b < c \;\&\; a \circ c < d \;.\!\!\rightarrow\; a \circ b < d$	[transitivity of $<$]
(4)	$a \circ b \not< c \;\&\; a \circ c \not< d \;.\!\!\rightarrow\; a \circ b \not< d$	[transitivity of $\not<$]

More generally, we can define an "anadic" predicate with finitely-many circle-operators as follows.

$a_1 \circ \ldots \circ a_k < b$ means $a_1, \ldots, a_k$ resemble each other more than any one of them resembles c

It satisfies basically the same principles concerning $<$ and $\not<$, and it satisfies an appropriate generalization of the symmetry principle, that allows arbitrary permutation of the factors in $a_1 \circ \ldots \circ a_k$.

22. Hierarchical Classification Systems

So far we have discussed simple classification, which may be distinguished from hierarchical classification. Whereas a simple classification divides a domain into one set $\{K_1, \ldots, K_m\}$ of categories, placing each item in exactly one category, a hierarchical classification places each item in a ranked series of categories.

This classification scheme gives rise to the notion of *species-genus*, which is a relation among categories and/or concepts. These two terms also have further important technical uses inside biology, for particular taxonomic ranks, which we discuss below, but we should keep in mind that these words (or their ancestors) were in common philosophical use thousands of years before they were appropriated by biologists. The basic idea is that one concept/category/kind K_1 can be a species of another concept/category/kind K_2. Furthermore, when K_1 is a species of K_2, the converse relation is that K_2 is a genus of K_1. Alternatively stated, K_1 is specific to K_2, and K_2 is generic to K_1.

A simple everyday example is that scarlet is a species of red, which is a species of color. In physics, a proton is a species of baryon, which is a species of elementary particle. In biology, humans are a species of hominid, which is a species of primate, which is a species of mammal. In linguistics, English is a species of Germanic language, which is a species of Indo-European language, which is a species of language.

Since Aristotle, it is has been customary to think of the species-genus relation in terms of *genus-cum-differentia*. According to this idea, a specific kind includes all the characteristics of the generic kind *plus* at least one supplementary distinctive (differentiating) characteristic. One of Aristotle's examples characterizes humans as rational animals; humans are animals distinguished by their rationality. This scheme is reflected in the modern binomial system of taxonomic nomenclature, originally due to Linnaeus,[45] according to which a biological species is designated by a pair of words – one for the genus, and one for the differentia (species). In many cases, the second name is an adjective, as in *Homo sapiens* (sapient "man").[46]

Note that the species-genus relation, which we may designate by '$\prec$', is transitive and asymmetric; i.e.:

---

[45] Alias Karl Linné (1707-1778), a Swedish botanist, whose Latinized name was 'Linnaeus'.

[46] It is interesting to note that the prefix 'homo' has two derivations, and accordingly two separate meanings. On the one hand, in words like 'homonym' and 'homogeneous', the prefix 'homo' derives from Greek *homos*, which means 'same'. On the other hand, in the species name *Homo sapiens*, the word derives from Latin *hom½*, which means 'man'. Notice that both the French word 'homme' [man] and the Spanish word 'hombre' [man] derive from this.

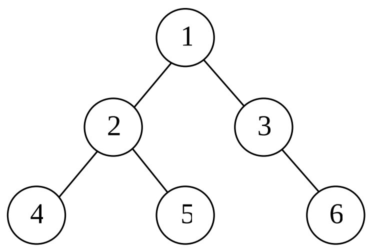
(t1) $x \prec y \ \& \ y \prec z \ \ .\!\rightarrow \ x \prec z$
(t2) $x \prec y \ \rightarrow \ y \nprec x$

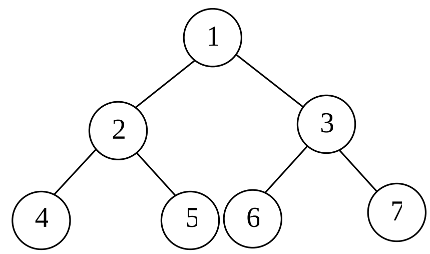
However, it need not be a linear-ordering, since its negation is not generally transitive. For example, humans are not a species of arthropod, and arthropods are not a species of mammal, but humans are a species of mammal. On the other hand, the relation $\prec$ satisfies the following alternative condition.

(t3) $x \prec y \ \& \ x \prec z \ \& \ y \neq z \ \ .\!\rightarrow\!. \ \ y \prec z \lor z \prec y$

In other words, if x is a species of two categories y and z, then one of these categories must be a species of the other. It follows that two categories cannot overlap unless one is a species of the other. Conditions (t1)-(t3) together characterize what mathematicians call a tree-ordering, or simply a "tree".

The following are two very simple examples of tree structures.





In these diagrams, a line connects two categories if one *directly* subsumes the other, the idea being that the higher category subsumes the lower category. For example, in the first diagram, the top-most category (*summum genus*) ① directly subsumes two categories ② and ③; then ② directly subsumes ④ and ⑤, and ③ directly subsumes only ⑥. Alternatively stated, ④ is a species of ②, which is a species of ①; ⑤ is also a species of ②. Also, ⑥ is a species of ③, which is a species of ①.

Another way to interpret tree structures is as family trees. For example, if the above diagrams represent patriarchal lineages (male family trees), then the first graph represents a man with two sons, one of which has two sons, and the other of which has one son, and the second graph represents a man with two sons, each of which has two sons. Notice that these trees branch downward, unlike real trees, but that is mathematically irrelevant. The main criterion that identifies them as trees is that branching occurs in only one direction. For example, a man can have numerous sons (downward branching), but only one father (no upward branching).

Probably the best-known classification system is the Linnaean classification system for biological organisms, which employs the following seven basic levels of organization as a starting point.

(1) kingdom
(2) phylum[47]
(3) class
(4) order
(5) family
(6) genus
(7) species

What makes a hierarchical classification system work is that, at each rank, an organism is placed in exactly one category – exactly one species, exactly one genus, and so forth. For example, humans are classified as follows, with black widow spiders thrown in for comparison.

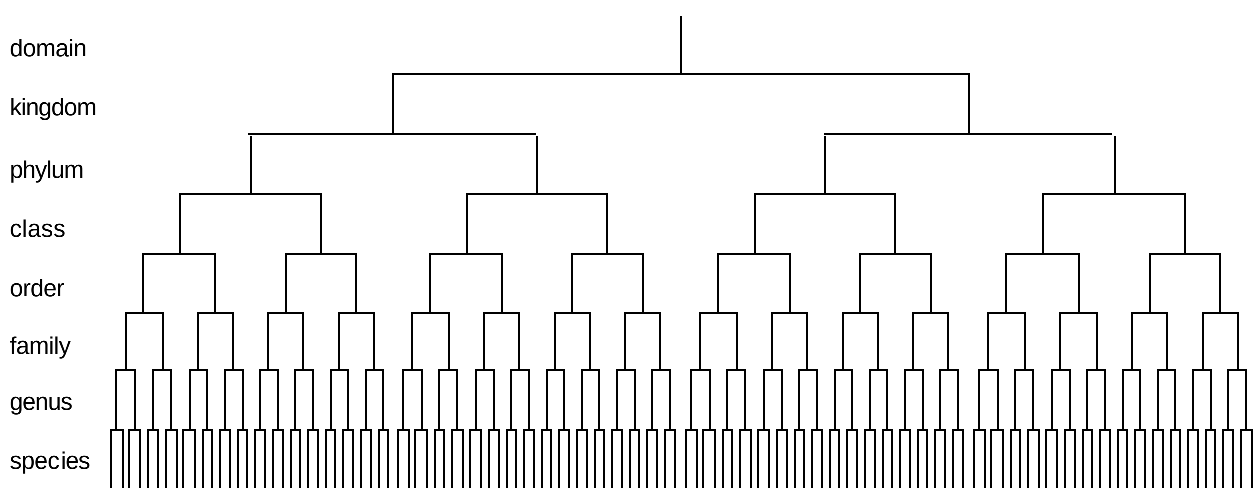
[47] Note that, in botany, the term ‘phylum’ is replaced by the term ‘division’. It’s a matter of tradition.

Linnaean Rank	**Humans**	**Black widows**
kingdom	Animals	Animals
phylum	Chordates	Arthropods
class	Mammals	Arachnids
order	Primates	Araneae
family	Hominids	Therididae
genus	*Homo*	*Latrodectus*
species	*sapiens*	*mactans*

Taxonomic classification systems can be viewed either top-down or bottom up. Viewed top-down, we imagine first dividing the general domain into very broad categories. For example, Aristotle divided the natural world into three broad categories – animate, vegetative, and inanimate. The modern counterpart of this is the folk-classification into animal, vegetable, and mineral. Currently, biologists divide living organisms into six kingdoms – animals, plants, fungi, protists, bacteria, and archae-bacteria.[48] Each kingdom is then divided into various phyla. For example, Aristotle divided animals according to their habitat – air, sea, land. On the other hand, his student Theophrastus divided plants according to their structure (number of stems) producing three categories – trees, shrubs, and herbs. In any case, each phylum is divided into classes, each of which is divided into families, each of which is divided into orders, each of which is divided into genera, each of which is divided into species.

Viewed bottom-up, we begin with organisms, and group them into species, which we group into genera, which we group into families, which we group into orders, which we group into classes, which we group into phyla, which we group into kingdoms.

Irrespective of whether we look at it top-down or bottom-up, we obtain a "tree of life". Viewed as a tree, the Linnaean hierarchy looks *generically* as follows, where we imagine that each category divides into two sub-categories.



[48] This classification has recently been called into question. The protists do not seem to form a single kingdom anymore. Unfortunately, these findings also undermine many of the other categories as well!