

Students raise this question so frequently that it might be best to try to deal with it in general terms: should **data** be used as a singular or as a plural noun? The historical answer is clear: the Latin **datum** is singular and **data** is its plural. The historical answer, however, may not be a good guide to current usage. The word **agenda**, for example, was originally plural (from **agendum**: 'something to be acted on') but is nowadays used only as a singular, and nobody in their right mind would insist that it should be used as a plural.

The following table summarises the results of an investigation of 750 randomly-selected citations for **data** from each of the following publications:

- The prestigious high-level research journal **Nature**.
- The weekly science magazine **New Scientist**
- The **Guardian** newspaper

The table shows the frequency of **data** in each publication: the proportion of citations in which **data** is unambiguously marked as plural or singular: and, of those, the proportion in which it is marked as plural and the proportion in which it is marked as singular:

| Publication | Frequency | Marked plur/sing | Plural | Singular |
|---------------|-----------|------------------|-------------|------------|
| Nature | 1:563 | 187 (23.6%) | 175 (93.6%) | 12 (6.4%) |
| New Scientist | 1:1,970 | 123 (16.4%) | 84 (69.9%) | 39 (30.1%) |
| Guardian | 1:22,047 | 81 (10.8%) | 32 (39.5%) | 49 (60.5%) |

The table shows that:

1. In 'high-level' writing on science (**Nature**), the use of **data** as a plural is dominant: in more popular writing on science (**New Scientist**) its use as a singular form is more frequent, though it is still outnumbered by its use as a plural: and in educated everyday usage as represented by the **Guardian** newspaper, it is nowadays most often used as a singular.
2. The difference between the use of **data** as a plural or singular correlates with the frequency with which the word is used (it is 40 times more frequent in **Nature**, and over 10 times more frequent in **New Scientist**, than in the **Guardian**), and also with the proportion of instances that are clearly marked as plural or singular (the highest proportion being in **Nature** and the lowest in the **Guardian**, with **New Scientist** again occupying an intermediate position).
3. The choice of plural or singular (and thus publication, frequency, and marking) also correlates with the **meaning** of the word. As the citations below (all taken from **New Scientist**) indicate, the (traditional) meaning "evidence used in experimental procedures" is most often plural, while the (more recent) meaning "digital information stored or manipulated by a computer" is most often singular. The best advice I can give is that if you are using the word in the first meaning, you should use it as a plural, but that you are probably quite safe in using it as a singular if it has the second meaning.

1 orresponding to its importance. The weighted data are then processed to produce a result. The
2 n of the experiment contained the gene. "Our data suggest that distances from engineered crop
3 What is perhaps more interesting is that the data appear to be independent of peer group. The
4 ghts into the way HIV might spread even when data on sexual behaviour are scarce. The first s
5 mber of animals taken in past seasons. Those data go into an assessment process in which the
6 analysis of the registry by the NRPB. These data show a correlation between lifetime exposur
7 tain about which treatments to use. But when data from large numbers of women are systematica
8 1989 and 1990, the last two years for which data are published, British discharges of heavy
9 sow the benefit wears off over time, but the data are "blind" so no one knows except the inde
10 known as Kornner returns. To cope with these data, hospitals bought large mainframe computers
11 ecord their morning round of checks. All the data is then dumped into the main computer. "If
12 normal circumstances, only half the digital data in the store is read. If the player is jogg
13 or the Inmarsat A satellites. The compressed data is fed onto the hard disc of a powerful per
14 otates at twice the normal speed so that the data is read off the disc twice as fast. All the
15 an just charitable feelings," says Dubnick. "Data is being digitised because it's easier to a
16 use, for a given image quality, it used less data than DCT-based systems. 'The data rate with
17 possible for a long time, if at all". Video data runs at a very high rate, too high for voic
18 nual layers downwards from the surface. "Our data strongly indicates that the last glaciation
19 ine research, concluding that "the available data provides no basis for testing any experimen
20 nuclear engineers also like to back up this data by examining samples from the pressure vess

Note that when it is used as a plural the word **data** cannot be preceded by a numeral: so instead of **"There are three data which support this hypothesis" write "There are three sets of data which support this hypothesis" or (depending on the situation) "There are three pieces of evidence which support this hypothesis."

March 1997 footnote. It seems that *criteria* (plural of *criterion*) has begun to journey along the path completed by *agenda* and on which *data* is already well under way. Of 1,504 citations from the Guardian for 1994 and Times for 1995, 14 (1%) are marked as being singular:

1. A criteria which measures the value of shopping, cleaning, cooking, caring for and maintaining oneself and one's intimates in terms of its economic contribution, for example, can find no grounds to contest the replacement of houseworkers and lovers by domestic servants and prostitutes.
2. Ask the pilots if you want a criteria for failure.
3. Detailed criteria **has** been set out as to what is meant by competence.
4. But optimists should still believe that ``by **every** criteria - from compassion to self-interest - a reduction in inequality is the sensible choice."
5. Our main criteria **is** to create products based on or inspired by Trust property.
6. The target they have to meet was reasoned out in 1957, and is known as the Lawson criteria. **It** is the formula conceived by J D Lawson, a British physicist, for calculating the combination of temperature, plasma density and time needed to sustain a fusion reaction.
7. Most relied on parents'; geographical proximity to the school as at least **one** criteria, despite the government's attempts to encourage parents to shop around, according to the study of 3,000 parents across 10 local authorities.
8. Using 'key negotiators' in each relevant regional office, Uni-key seeks out properties which fit the strict guidelines set out by universities (within a bike ride of the campus is **one** criteria).
9. What other criteria **is** there besides ability?
10. Unless the criteria **is** published then the explanation remains a confidence trick.
11. "In the odd place where creativity meets commerce, which is Hollywood among other places, their criteria **is** very simple: to make money.
12. The system has already tuned in to this criteria.
13. MHEP fits **this** criteria.
14. While **safety** must always be the ultimate criteria, there is no reason to ignore technological improvements made since Taylor reported.

It will be interesting over the next 10-20 years to see whether *criteria* continues on the route towards singularity.

[More recent evidence about *data* and *criteria*.](#)

16th June 1996 and 3rd March 1997

[Back to Kibbitzers](#)

Consultant: Tim Johns