

3. The objectivity of science

- 3.1 Data is the lifeblood of scientists and may be qualitative or quantitative. It can be obtained purely from observations or from specifically designed experiments, remotely using electronic sensors or by direct measurement. The best data for making accurate and precise descriptions and predictions is often quantitative and amenable to mathematical analysis. Scientists analyze data and look for patterns, trends and discrepancies, attempting to discover relationships and establish causal links. This is not always possible, so identifying and classifying observations and artefacts (e.g. types of galaxies or fossils) is still an important aspect of scientific work.
- 3.2 Taking repeated measurements and large numbers of readings can improve reliability in data collection. Data can be presented in a variety of formats such as linear and logarithmic graphs that can be analyzed for, say, direct or inverse proportion or for power relationships.
- 3.3 Scientists need to be aware of random errors and systematic errors, and use techniques such as error bars and lines of best fit on graphs to portray the data as realistically and honestly as possible. There is a need to consider whether outlying data points should be discarded or not.
- 3.4 Scientists need to understand the difference between errors and uncertainties, accuracy and precision, and need to understand and use the mathematical ideas of average, mean, mode, median, etc. Statistical methods such as standard deviation and chi-squared tests are often used. It is important to be able to assess how accurate a result is. A key part of the training and skill of scientists is in being able to decide which technique is appropriate in different circumstances.
- 3.5 It is also very important for scientists to be aware of the cognitive biases that may impact experimental design and interpretation. The confirmation bias, for example, is a well-documented cognitive bias that urges us to find reasons to reject data that is unexpected or does not conform to our expectations or desires, and to perhaps too readily accept data that agrees with these expectations or desires. The processes and methodologies of science are largely designed to account for these biases. However, care must always be taken to avoid succumbing to them.
- 3.6 Although scientists cannot ever be certain that a result or finding is correct, we know that some scientific results are very close to certainty. Scientists often speak of “levels of confidence” when discussing outcomes. The discovery of the existence of a Higgs boson is such an example of a “level of confidence”. This particle may never be directly observable, but to establish its “existence” particle physicists had to pass the self-imposed definition of what can be regarded as a discovery—the 5-sigma “level of certainty”—or about a 0.00003% chance that the effect is not real based on experimental evidence.

- 3.7 In recent decades, the growth in computing power, sensor technology and networks has allowed scientists to collect large amounts of data. Streams of data are downloaded continuously from many sources such as remote sensing satellites and space probes and large amounts of data are generated in gene sequencing machines. Experiments in CERN's Large Hadron Collider regularly produce 23 petabytes of data per second, which is equivalent to 13.3 years of high definition TV content per second.
- 3.8 Research involves analyzing large amounts of this data, stored in databases, looking for patterns and unique events. This has to be done using software which is generally written by the scientists involved. The data and the software may not be published with the scientific results but would be made generally available to other researchers.