

4. Analyzing and Interpreting Data

Scientific investigations produce data that must be analyzed in order to derive meaning. Because data usually do not speak for themselves, scientists use a range of tools—including tabulation, graphical interpretation, visualization, and statistical analysis—to identify the significant features and patterns in the data. Sources of error are identified and the degree of certainty calculated. Modern technology makes the collection of large data sets much easier, thus providing many secondary sources for analysis.

Engineers analyze data collected in the tests of their designs and investigations; this allows them to compare different solutions and determine how well each one meets specific design criteria—that is, which design best solves the problem within the given constraints. Like scientists, engineers require a range of tools to identify the major patterns and interpret the results.

5. Using Mathematics and Computational Thinking

In **science**, mathematics and computation are fundamental tools for representing physical variables and their relationships. They are used for a range of tasks, such as constructing simulations, statistically analyzing data, and recognizing, expressing, and applying quantitative relationships. Mathematical and computational approaches enable predictions of the behavior of physical systems, along with the testing of such predictions. Moreover, statistical techniques are invaluable for assessing the significance of patterns or correlations.

In **engineering**, mathematical and computational representations of established relationships and principles are an integral part of design. For example, structural engineers create mathematically based analyses of designs to calculate whether they can stand up to the expected stresses of use and if they can be completed within acceptable budgets. Moreover, simulations of designs provide an effective test bed for the development of designs and their improvement.

opportunities to carry out careful and systematic investigations, with appropriately supported prior experiences that develop their ability to observe and measure and to record data using appropriate tools and instruments.

Students should have opportunities to plan and carry out several different kinds of investigations during their K-12 years. At all levels, they should engage in investigations that range from those structured by the teacher—in order to expose an issue or question that they would be unlikely to explore on their own (e.g., measuring specific properties of materials)—to those that emerge from students' own



questions. As they become more sophisticated, students also should have opportunities not only to identify questions to be researched but also to decide what data are to be gathered, what variables should be controlled, what tools or instruments are needed to gather and record data in an appropriate format, and eventually to consider how to incorporate measurement error in analyzing data.

Older students should be asked to develop a hypothesis that predicts a particular and stable outcome and to explain their reasoning and justify their choice. By high school, any hypothesis should be based on a well-developed model or theory. In addition, students should be able to recognize that it is not always possible to control variables and that other methods can be used in such cases—for example, looking for correlations (with the understanding that correlations do not necessarily imply causality).

Practice 4 Analyzing and Interpreting Data

Once collected, data must be presented in a form that can reveal any patterns and relationships and that allows results to be communicated to others. Because raw data as such have little meaning, a major practice of scientists is to organize and interpret data through tabulating, graphing, or statistical analysis. Such analysis can bring out the meaning of data—and their relevance—so that they may be used as evidence.

■ Once collected, data must be presented in a form that can reveal any patterns and relationships and that allows results to be communicated to others. ■

Engineers, too, make decisions based on evidence that a given design will work; they rarely rely on trial and error. Engineers often analyze a design by creating a model or prototype and collecting extensive data on how it performs, including under extreme conditions. Analysis of this kind of data not only informs design decisions and enables the prediction or assessment of performance but also helps define or clarify problems, determine economic feasibility, evaluate alternatives, and investigate failures.

Spreadsheets and databases provide useful ways of organizing data, especially large data sets. The identification of relationships in data is aided by a range of tools, including tables, graphs, and mathematics. Tables permit major features of a large body of data to be summarized in a conveniently accessible form, graphs offer a means of visually summarizing data, and mathematics is essential for expressing relationships between different variables in the data set (see Practice 5 for further discussion of mathematics). Modern computer-based visualization tools often allow data to be displayed in varied forms and thus for learners to engage interactively with data in their analyses. In addition, standard statistical techniques can help to reduce the effect of error in relating one variable to another.

Students need opportunities to analyze large data sets and identify correlations. Increasingly, such data sets—involving temperature, pollution levels, and other scientific measurements—are available on the Internet. Moreover, information technology enables the capture of data beyond the classroom at all hours of the day. Such data sets extend the range of students' experiences and help to illuminate this important practice of analyzing and interpreting data.

GOALS

By grade 12, students should be able to

- Analyze data systematically, either to look for salient patterns or to test whether data are consistent with an initial hypothesis.
- Recognize when data are in conflict with expectations and consider what revisions in the initial model are needed.

- Use spreadsheets, databases, tables, charts, graphs, statistics, mathematics, and information and computer technology to collate, summarize, and display data and to explore relationships between variables, especially those representing input and output.
- Evaluate the strength of a conclusion that can be inferred from any data set, using appropriate grade-level mathematical and statistical techniques.
- Recognize patterns in data that suggest relationships worth investigating further. Distinguish between causal and correlational relationships.
- Collect data from physical models and analyze the performance of a design under a range of conditions.

PROGRESSION

At the elementary level, students need support to recognize the need to record observations—whether in drawings, words, or numbers—and to share them with others. As they engage in scientific inquiry more deeply, they should begin to collect categorical or numerical data for presentation in forms that facilitate interpretation, such as tables and graphs. When feasible, computers and other digital tools should be introduced as a means of enabling this practice.

In middle school, students should have opportunities to learn standard techniques for displaying, analyzing, and interpreting data; such techniques include different types of graphs, the identification of outliers in the data set, and averaging to reduce the effects of measurement error. Students should also be asked to explain why these techniques are needed.

As students progress through various science classes in high school and their investigations become more complex, they need to develop skill in additional techniques for displaying and analyzing data, such as x-y scatterplots or cross-tabulations to express the relationship between two variables. Students should be helped to recognize that they may need to explore more than one way to display their data in order to identify and present significant features. They also need opportunities to use mathematics and statistics to analyze features of data such as covariation. Also at the high school level, students should have the opportunity to use a greater diversity of samples of scientific data and to use computers or other digital tools to support this kind of analysis.

Students should be expected to use some of these same techniques in engineering as well. When they do so, it is important that they are made cognizant of the purpose of the exercise—that any data they collect and analyze are intended to help validate or improve a design or decide on an optimal solution.