

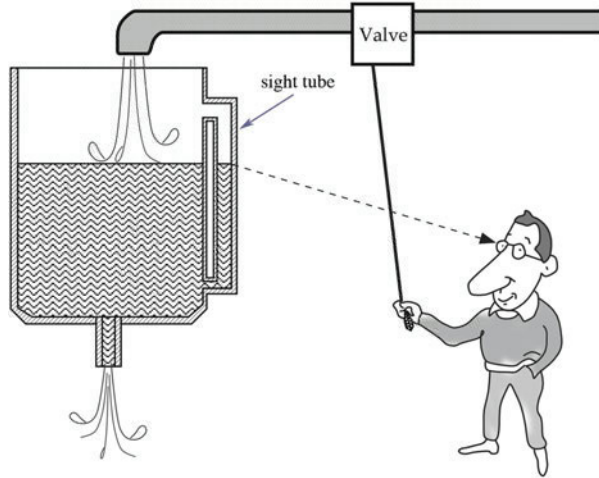
“It’s as large as life, and twice as natural”

—Lewis Carroll, “Through the Looking Glass”

1.1 Sensors, Signals, and Systems

A sensor is often defined as a “*device that receives and responds to a signal or stimulus*”. This definition is broad. In fact, it is so broad that it covers almost everything from a human eye to a trigger in a pistol. Consider the level-control system shown in Fig. 1.1 [1]. The operator adjusts the level of fluid in the tank by manipulating its valve. Variations in the inlet flow rate, temperature changes (these would alter the fluid’s viscosity and consequently the flow rate through the valve), and similar disturbances must be compensated for by the operator. Without control the tank is likely to flood, or run dry. To act appropriately, the operator must on a timely basis obtain information about the level of fluid in the tank. In this example, the information is generated by the sensor, which consists of two main parts: the sight tube on the tank and the operator’s eye, which produces an electric response in the optic nerve. The sight tube by itself is not a sensor, and in this particular control system, the eye is not a sensor either. Only the combination of these two components makes a narrow-purpose sensor (detector) that is *selectively* sensitive to the fluid level. If a sight tube is designed properly, it will very quickly reflect variations in the level, and it is said that the sensor has a fast speed response. If the internal diameter of the tube is too small for a given fluid viscosity, the level in the tube may lag behind the level in the tank. Then, we have to consider a phase characteristic of such a sensor. In some cases, the lag may be quite acceptable, while in other situations, a better sight tube design would be required. Hence, the sensor’s performance must be assessed only as part of a data acquisition system.

Fig. 1.1 Level-Control System. Sight tube and operator's eye form a sensor—device that converts information into electrical signal



This world is divided into natural and man-made objects. The natural sensors, like those found in living organisms, usually respond with signals having electrochemical character; that is, their physical nature is based on ion transport, like in the nerve fibers (such as an optic nerve in the fluid tank operator). In man-made devices, information is also transmitted and processed in electrical form, however, through the transport of electrons. Sensors intended for the artificial systems must speak the same language as the systems “speak”. This language is electrical in its nature and the sensor shall be capable of responding with the output signals where information is carried by displacement of electrons, rather than ions.¹ Thus, it should be possible to connect a sensor to an electronic system through electrical wires, rather than through an electrochemical solution or a nerve fiber. Hence, in this book, we use a somewhat narrower definition of a sensor, which may be phrased as

A sensor is a device that receives a stimulus and responds with an electrical signal.

The term *stimulus* is used throughout this book and needs to be clearly understood. The stimulus is the quantity, property, or condition that is received and converted into electrical signal. Examples of stimuli are light intensity and wavelength, sound, force, acceleration, distance, rate of motion, and chemical composition. When we say “electrical,” we mean a signal which can be channeled, amplified, and modified by electronic devices. Some texts (for instance, [2]) use a different term, *measurand*, which has the same meaning as stimulus, however with the stress on quantitative characteristic of sensing.

We may say that a sensor is a translator of a generally nonelectrical value into an electrical value. The sensor's output signal may be in form of voltage, current, or charge. These may be further described in terms of amplitude, polarity, frequency,

¹ There is a very exciting field of the optical computing and communications where information is processed by a transport of photons. That field is beyond the scope of this book.

phase, or digital code. The set of output characteristics is called the *output signal format*. Therefore, a sensor has input properties (of any kind) and electrical output properties.

Any sensor is an energy converter. No matter what you try to measure, you always deal with energy transfer between the object of measurement to the sensor. The process of sensing is a particular case of information transfer, and any transmission of information requires transmission of energy. One should not be confused by the obvious fact that transmission of energy can flow both ways—it may be with a positive sign as well as with a negative sign; that is, energy can flow either from the object to the sensor or backward—from the sensor to the object. A special case is when the net energy flow is zero, and that also carries information about existence of that particular situation. For example, a thermopile infrared radiation sensor will produce a positive voltage when the object is warmer than the sensor (infrared flux is flowing to the sensor). The voltage becomes negative when the object is cooler than the sensor (infrared flux flows from the sensor to the object). When both the sensor and the object are at exactly the same temperature, the flux is zero and the output voltage is zero. This carries a message that the temperatures are equal to one another.

The terms *sensor* and term *detector* are synonyms, used interchangeably and have the same meaning. However, detector is more often used to stress qualitative rather than quantitative nature of measurement. For example, a PIR (passive infrared) detector is employed to indicate just the existence of human movement but generally cannot measure direction, speed, or acceleration.

The term *sensor* should be distinguished from *transducer*. The latter is a converter of any one type of energy or property into another type of energy or property, whereas the former converts it into *electrical signal*. An example of a transducer is a loudspeaker which converts an electrical signal into a variable magnetic field and, subsequently, into acoustic waves.² This is nothing to do with perception or sensing. Transducers may be used as *actuators* in various systems. An actuator may be described as opposite to a sensor—it converts electrical signal into generally nonelectrical energy. For example, an electric motor is an actuator—it converts electric energy into mechanical action. Another example is a pneumatic actuator that is enabled by an electric signal and converts air pressure into force.

Transducers may be parts of a *hybrid* or *complex* sensor (Fig. 1.2). For example, a chemical sensor may comprise two parts: the first part converts energy of an exothermal chemical reaction into heat (transducer) and another part, a thermopile, converts heat into an electrical output signal. The combination of the two makes a hybrid chemical sensor, a device which produces *electrical* signal in response to a chemical reagent. Note that in the above example a chemical sensor is a complex sensor—it is comprised of a nonelectrical transducer and a simple (direct) sensor converting heat to electricity. This suggests that many sensors incorporate at least

²It is interesting to note that a loudspeaker, when connected to an input of an amplifier, may function as a microphone. In that case, it becomes an acoustical sensor.

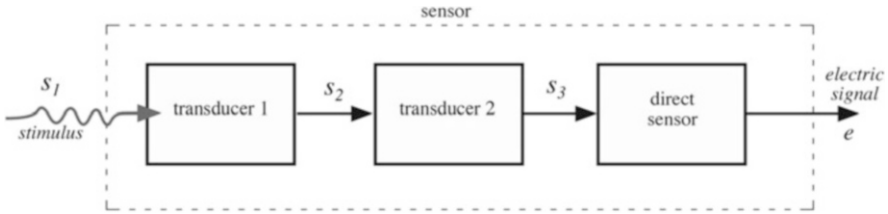


Fig. 1.2 Sensor may incorporate several transducers. Value s_1 , s_2 , etc. represent various types of energy. Direct sensor produces electrical output e

one *direct*-type sensor and possibly a number of transducers. The direct sensors are those that employ certain physical effects to make a *direct* energy conversion into a generation or modulation of an electrical signal. Examples of such physical effects are the photoeffect and Seebeck effect. These will be described in Chap. 4.

In summary, there are two types of sensors, *direct* and *hybrid*. A direct sensor converts a stimulus into an electrical signal or modifies an externally supplied electrical signal, whereas a hybrid sensor (or simply—a sensor) in addition needs one or more transducers before a direct sensor can be employed to generate an electrical output.

A sensor does not function by itself; it is always part of a larger system that may incorporate many other detectors, signal conditioners, processors, memory devices, data recorders, and actuators. The sensor's place in a device is either intrinsic or extrinsic. It may be positioned at the input of a device to perceive the outside effects and to inform the system about variations in the outside stimuli. Also, it may be an internal part of a device that monitors the devices' own state to cause the appropriate performance. A sensor is always part of some kind of a data acquisition system. In turn, such a system may be part of a larger control system that includes various feedback mechanisms.

To illustrate the place of sensors in a larger system, Fig. 1.3 shows a block diagram of a data acquisition and control device. An object can be anything: a car, space ship, animal or human, liquid, or gas. Any material object may become a subject of some kind of a measurement or control. Data are collected from an object by a number of sensors. Some of them (2, 3, and 4) are positioned directly on or inside the object. Sensor 1 perceives the object without a physical contact and, therefore, is called a *noncontact* sensor. Examples of such a sensor is a radiation detector and a TV camera. Even if we say “noncontact”, we remember that energy transfer always occurs between a sensor and object.

Sensor 5 serves a different purpose. It monitors the internal conditions of the data acquisition system itself. Some sensors (1 and 3) cannot be directly connected to standard electronic circuits because of the inappropriate output signal formats. They require the use of interface devices (signal conditioners) to produce a specific output format.

Sensors 1, 2, 3, and 5 are *passive*. They generate electric signals without energy consumption from the electronic circuits. Sensor 4 is *active*. It requires an operating

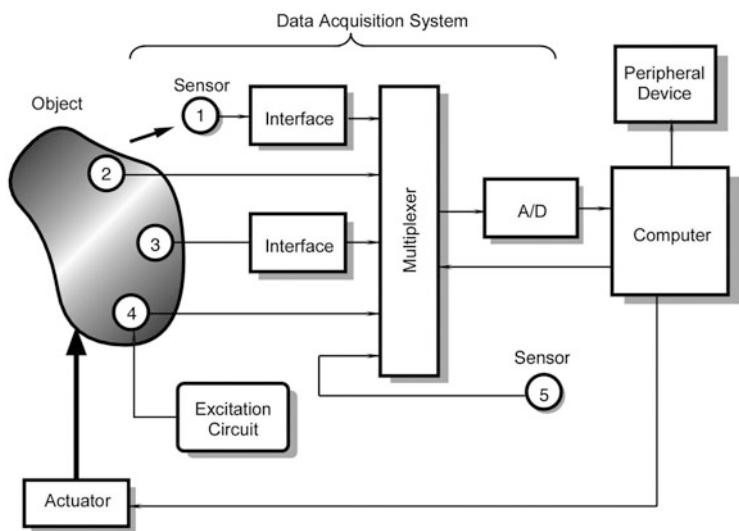


Fig. 1.3 Positions of sensors in data acquisition system. Sensor 1 is noncontact, sensors, 2 and 3 are passive, sensor 4 is active, and sensor 5 is internal to data acquisition system

signal that is provided by an excitation circuit. This signal is modified by the sensor or modulated by the object’s stimulus. An example of an active sensor is a thermistor that is a temperature-sensitive resistor. It needs a current source, which is an excitation circuit. Depending on the complexity of the system, the total number of sensors may vary from as little as one (a home thermostat) to many thousands (a space station).

Electrical signals from the sensors are fed into a multiplexer (MUX), which is a switch or a gate. Its function is to connect the sensors, one at a time, to an analog-to-digital converter (A/D or ADC) if a sensor produces an analog signal, or directly to a computer if a sensor produces signals in a digital format. The computer controls a multiplexer and ADC for the appropriate timing. Also, it may send control signals to an actuator that acts on the object. Examples of the actuators are an electric motor, a solenoid, a relay, and a pneumatic valve. The system contains some peripheral devices (for instance, a data recorder, display, alarm, etc.) and a number of components that are not shown in the block diagram. These may be filters, sample-and-hold circuits, amplifiers, and so forth.

To illustrate how such a system works, let us consider a simple car door monitoring arrangement. Every door in a car is supplied with a sensor that detects the door position (open or closed). In most cars, the sensor is a simple electric switch. Signals from all door switches go to the car’s internal processor (no need for an ADC as all door signals are in a digital format: ones or zeros). The processor identifies which door is open (signal is zero) and sends an indicating message to the peripheral devices (a dashboard display and an audible alarm). A car driver (the actuator) gets the message and acts on the object (closes the door) and the sensor outputs the signal “one”.

An example of a more complex device is an anesthetic vapor delivery system. It is intended for controlling the level of anesthetic drugs delivered to a patient through inhalation during surgical procedures. The system employs several active and passive sensors. The vapor concentration of anesthetic agents (such as halothane, isoflurane, or enflurane) is selectively monitored by an active piezoelectric sensor being installed into a ventilation tube. Molecules of anesthetic vapors add mass to the oscillating crystal in the sensor and change its natural frequency, which is a measure of the vapor concentration. Several other sensors monitor the concentration of CO_2 , to distinguish exhale from inhale, and temperature and pressure, to compensate for additional variables. All these data are multiplexed, digitized, and fed into the digital signal processor (DSP) which calculates the actual vapor concentration. An anesthesiologist presets a desired delivery level and the processor adjusts the actuators (valves) to maintain anesthetics at the correct concentration.

Another example of a complex combination of various sensors, actuators, and indicating signals is shown in Fig. 1.4. It is an Advanced Safety Vehicle (ASV) that was developed by Nissan. The system is aimed at increasing safety of a car. Among many others, it includes a drowsiness warning system and drowsiness relieving system. This may include the eyeball movement sensor and the driver head inclination detector. The microwave, ultrasonic, and infrared range measuring sensors are incorporated into the emergency braking advanced advisory system to illuminate the break lamps even before the driver brakes hard in an emergency, thus advising the driver of a following vehicle to take evasive action. The obstacle warning system includes both the radar and infrared (IR) detectors. The adaptive cruise-control system works if the driver approaches too closely to a preceding vehicle; the speed is automatically reduced to maintain a suitable safety distance. The pedestrian monitoring system detects and alerts the driver to the presence of pedestrians at night as well as in vehicle blind spots. The lane-control system helps in the event the system detects and determines that incipient lane deviation is not the driver's intention. It issues a warning and automatically steers the vehicle, if necessary, to prevent it from leaving its lane.

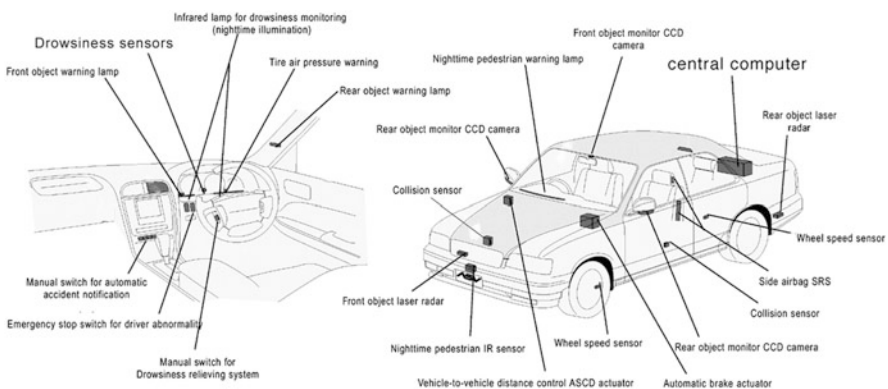


Fig. 1.4 Multiple sensors, actuators, and warning signals are parts of the Advanced Safety Vehicle (Courtesy of Nissan Motor Company)

In the following chapters we focus on sensing methods, physical principles of sensor operations, practical designs, and interface electronic circuits. Other essential parts of the control and monitoring systems, such as actuators, displays, data recorders, data transmitters, and others are beyond the scope of this book and mentioned only briefly.

The sensor's packaging design may be of a general purpose. A special packaging and housing should be built to adapt it for a particular application. For instance, a micromachined piezoresistive pressure sensor may be housed into a watertight enclosure for the invasive measurement of the aortic blood pressure through a catheter. The same sensor will be given an entirely different packaging when intended for measuring blood pressure by a noninvasive oscillometric method with an inflatable cuff. Some sensors are specifically designed to be very selective in a particular range of input stimulus and be quite immune to signals outside the desirable limits. For instance, a motion detector for a security system should be sensitive to movement of humans and not responsive to movement of smaller animals, like dogs and cats.

1.2 Sensor Classification

Sensor classification schemes range from very simple to the complex. Depending on the classification purpose, different classification criteria may be selected. Here are several practical ways to look at sensors.

1. All sensors may be of two kinds: *passive* and *active*. A passive sensor does not need any additional energy source. It generates an electric signal in response to an external stimulus. That is, the input stimulus energy is converted by the sensor into the output signal. The examples are a thermocouple, a photodiode, and a piezoelectric sensor. Many passive sensors are *direct* sensors as we defined them earlier.

The *active* sensors require external power for their operation, which is called an *excitation signal*. That signal is modified (modulated) by the sensor to produce the output signal. The active sensors sometimes are called *parametric* because their own properties change in response to an external stimulus and these properties can be subsequently converted into electric signals. It can be stated that a sensor's parameter modulates the excitation signal and that modulation carries information of the measured value. For example, a thermistor is a temperature-sensitive resistor. It does not generate any electric signal, but by passing electric current (excitation signal) through it its resistance can be measured by detecting variations in current and/or voltage across the thermistor. These variations (presented in ohms) directly relate to temperature through a known transfer function. Another example of an active sensor is a resistive strain gauge in which electrical resistance relates to strain in the material. To measure the resistance of a sensor, electric current must be applied to it from an external power source.

2. Depending on the selected reference, sensors can be classified into *absolute* and *relative*. An absolute sensor detects a stimulus in reference to an absolute physical scale that is independent on the measurement conditions, whereas a relative sensor produces a signal that relates to some special case. An example of an absolute sensor is a thermistor—a temperature-sensitive resistor. Its electrical resistance directly relates to the absolute temperature scale of Kelvin. Another very popular temperature sensor—a thermocouple—is a relative sensor. It produces an electric voltage that is function of a temperature gradient across the thermocouple wires. Thus, a thermocouple output signal cannot be related to any particular temperature without referencing to a selected baseline. Another example of the absolute and relative sensors is a pressure sensor. An absolute pressure sensor produces signal in reference to vacuum—an absolute zero on a pressure scale. A relative pressure sensor produces signal with respect to a selected baseline that is not zero pressure—for example, to the atmospheric pressure.
3. Another way to look at a sensor is to consider some of its properties that may be of a specific interest [3]. Below are the lists of various sensor characteristics and properties (Tables 1.1, 1.2, 1.3, 1.4, and 1.5).

Table 1.1 Sensor specifications

Sensitivity	Stimulus range (span)
Stability (short and long term)	Resolution
Accuracy	Selectivity
Speed of response	Environmental conditions
Overload characteristics	Linearity
Hysteresis	Dead band
Operating life	Output format
Cost, size, weight	Other

Table 1.2 Sensing element material

Inorganic	Organic
Conductor	Insulator
Semiconductor	Liquid gas or plasma
Biological substance	Other

Table 1.3 Conversion phenomena

Physical	Thermoelectric Photoelectric Photomagnetic Magnetoelectric Electromagnetic Thermoelastic Electroelastic Thermomagnetic Thermo-optic Photoelastic Other	Chemical	Chemical transformation Physical transformation Electrochemical process Spectroscopy Other
		Biological	Biochemical transformation Physical transformation Effect on test organism Spectroscopy Other

Table 1.4 Field of applications

Agriculture	Automotive
Civil engineering, construction	Domestic, appliances
Distribution, commerce, finance	Environment, meteorology, security
Energy, power	Information, telecommunication
Health, medicine	Marine
Manufacturing	Recreation, toys
Military	Space
Scientific measurement	Other
Transportation (excluding automotive)	

Table 1.5 Stimuli

Stimulus	Stimulus	
<i>Acoustic</i> Wave amplitude, phase Spectrum polarization Wave velocity Other	<i>Mechanical</i>	Position (linear, angular) Acceleration Force Stress, pressure Strain
<i>Biological</i> Biomass (types, concentration states) Other		Mass, density Moment, torque Speed of flow, rate of mass transport
<i>Chemical</i> Components (identities, concentration, states) Other		Shape, roughness, orientation Stiffness, compliance
<i>Electric</i> Charge, current Potential, voltage Electric field (amplitude, phase, polarization, spectrum) Conductivity Permittivity Other		Viscosity Crystallinity, structural integrity Other
<i>Magnetic</i> Magnetic field (amplitude, phase, polarization, spectrum) Magnetic flux Permeability Other	<i>Radiation</i>	Type Energy Intensity Other
<i>Optical</i> Wave amplitude, phase, polarization, spectrum Wave velocity Refractive index Emissivity, reflectivity, absorption Other	<i>Thermal</i>	Temperature Flux Specific heat Thermal conductivity Other

1.3 Units of Measurements

In this book, we use base units which have been established in The 14th General Conference on Weights and Measures (1971). The base measurement system is known as SI which stands for French “*Le Système International d’Unités*” (Table 1.6) [4]. All other physical quantities are derivatives of these base units.³ Some of them are listed in Table A.3.

Often it is not convenient to use base or derivative units directly—in practice quantities may be either too large or too small. For convenience in the engineering work, multiples and submultiples of the units are generally employed. They can be obtained by multiplying a unit by a factor from the Appendix Table A.2. When pronounced, in all cases the first syllable is accented. For example, 1 ampere (A) may be multiplied by factor of 10^{-3} to obtain a smaller unit; 1 milliampere (1 mA) which is one thousandth of an ampere or 1 kilohm (1 kΩ) is one thousands of Ohms, where 1 Ω is multiplied by 10^3 .

Sometimes, two other systems of units are used. They are the Gaussian System and the British System, and in the U.S.A. its modification is called the

Table 1.6 SI basic units

Quantity	Name	Symbol	Defined by... (year established)
Length	meter	m	...the length of the path traveled by light in vacuum in 1/299,792,458 of a second. ... (1983)
Mass	kilogram	kg	...after a platinum-iridium prototype (1889)
Time	second	s	...the duration of 9,192,631,770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium-133 atom (1967)
Electric current	ampere	A	force equal to 2×10^{-7} N/m of length exerted on two parallel conductors in vacuum when they carry the current (1946)
Thermodynamic temperature	kelvin	K	The fraction 1/273.16 of the thermodynamic temperature of the triple point of water (1967)
Amount of substance	mole	mol	...the amount of substance which contains as many elementary entities as there are atoms in 0.012 kg of carbon 12 (1971)
Luminous intensity	candela	cd	...intensity in the perpendicular direction of a surface of 1/600,000 m ² of a blackbody at temperature of freezing Pt under pressure of 101,325 N/m ² (1967)
Plane angle	radian	rad	(supplemental unit)
Solid angle	steradian	sr	(supplemental unit)

³The SI is often called the *modernized metric system*.

US Customary System. The United States is the only developed country where SI still is not in common use. However, with the increase of globalization, it appears unavoidable that America will convert to SI in the future, though perhaps not in our lifetime. Still, in this book, we will generally use SI; however, for the convenience of the reader, the US customary system units will be used in places where US manufacturers employ them for the sensor specifications.

For conversion to SI from other systems⁴ use Table A.4 of the Appendix. To make a conversion, a non-SI value should be multiplied by a number given in the table. For instance, to convert acceleration of 55 ft/s² to SI, it must to be multiplied by 0.3048:

$$55 \text{ ft/s}^2 \times 0.3048 = 16.764 \text{ m/s}^2$$

Similarly, to convert electric charge of 1.7 faraday, it must be multiplied by 9.65×10^{19} :

$$1.7 \text{ faraday} \times 9.65 \times 10^{19} = 1.64 \times 10^{20} \text{ C}$$

The reader should consider a correct terminology of the physical and technical terms. For example, in the U.S.A. and several other countries, electric potential difference is called “*voltage*”, while in other countries “*electric tension*” or simply “*tension*” is in common use, such as *spannung* in German, *напряжение* in Russian, *tensione* in Italian, and 电压 in Chinese. In this book, we use terminology that is traditional in the United States of America.

References

1. Thompson, S. (1989). *Control systems: Engineering & design*. Essex, England: Longman Scientific & Technical.
2. Norton, H. N. (1989). *Handbook of transducers*. Englewood Cliffs, NJ: Prentice Hall.
3. White, R. W. (1991). A sensor classification scheme. In *Microsensors* (pp. 3–5). New York: IEEE Press.
4. Thompson, A., & Taylor, B. N. (2008). *Guide for the use of the international system of units (SI)*. NIST Special Publication 811, National Institute of Standards and Technology, Gaithersburg, MD 20899, March 2008.

⁴ Nomenclature, abbreviations, and spelling in the conversion tables are in accordance with ASTM SI10-02 IEEE/ASTM SI10 *American National Standard for Use of the International System of Units (SI): The Modern Metric System*. A copy is available from ASTM, 100 Barr Harbor Dr., West Conshocken, PA 19428-2959, USA. www.astm.org/Standards/SI10.htm