4 SPATIAL DISPLAYS

When we drive a car, we derive information about the depth and position of other objects in the world from the scene through the windshield. Similarly, when we examine a bar graph or check a speedometer, we derive information about the state of the world from a spatial array. The sizes of objects or the distances between them are used to communicate the relevant information. Human performance in such spatial judgments depends on accurate judgments of distance, extent, and depth. Our ability to perceive and understand such spatial relations will be the focus of this chapter.

Generally, large spatial or physical differences are more important or significant than small ones. Consider reading a graph or an analog meter. A small change in position reflects a small change in the underlying dimension. In contrast, consider reading a digital meter or a word. In a digital meter the spatial difference in the *physical* representation between, say, 79999 and 80000 is substantial—every digit is changed. But the difference in *meaning* between these two values is small. An analog display preserves some of the inherent properties of the dimension it represents: in this sense, it is an analog of its physical counterpart.

In this chapter, we consider a variety of spatial displays. We first discuss the perception and understanding of graphs. Then we address the role of motion as we consider the design of common displays such as meters and dials. In doing so, we highlight the importance of *compatibility* between the dimension portrayed and display elements. We consider compatibility in both static and dynamic senses. Space, of course, is also *three dimensional (3D)*. Our perception of a 3D environment is determined by the information we obtain about its structure as we move through it. We thus consider the various types of information we can obtain through movement, and their implication for display design. We are also concerned with perceptual judgment of depth and distance. We discuss the implications of such judgment on perception of real-world environments and for representing a 3D space on a 2D display surface. We close the chapter with a brief discussion of spatial displays that use other sensory modalities. In Chapter 5 we will expand on some of these topics while examining navigation and interaction with real and virtual environments.

1. GRAPH PERCEPTION

Unlike many of the displays discussed in this book, most of us will, at one time or another, design a graph. In the process, we must make decisions about graph type, assign variables to axes, code variables using symbols, and so on. This makes the graph a good place to start a discussion of display design. We define a *graph* as a paper or electronic representation of numeric analog data with multiple data points. Some everyday examples—bar graphs, line graphs, and pie charts—are shown in Figure 4.1. The distinction between graphs and analog displays has become blurred in recent years due to developments in information visualization, where graphs can dynamically change from one format to another, for example (Heer & Robertson, 2007; see Chapter 5), but one remaining difference is that with graphs, the data typically do not change as the user views

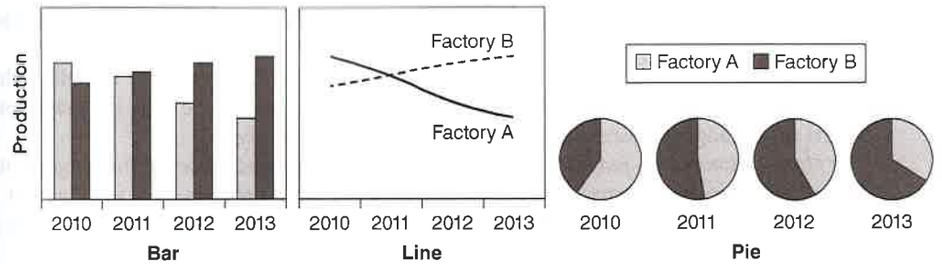

FIGURE 4.1 An example of a bar graph, a line graph, and a set of pie charts. Each graph type depicts the same data: the production of two factories, A and B, over four years. Four graph reading tasks that could be performed with each graph are also described.

them, whereas with information displays the data shown can change in real time as the user performs tasks and monitors the outcome.

A history of the graphic display of data dates back to the pioneering work of Playfair (1786), who first realized the power of using analog representations (e.g., bar graph, pie chart) to represent quantitative data. For spatial judgments (e.g., which variable is decreasing more quickly?), performance is better with graphs than tables (e.g., Kirschenbaum & Arruda, 1994; Vessey, 1991). As noted above, for spatial judgments large differences between values are more significant than small ones. It comes as no surprise, therefore, that an analog representation like a graph is more effective for the spatial judgment than a digital display. In contrast, reading a precise value is generally performed better with tables of digits (Lalomia, Coovert, & Salas, 1992; Meyer, Shinar, & Leiser, 1997; Vessey, 1991).

In Chapter 1, we introduced a model of human information processing. When considering the processing of graphs, we are looking primarily at the perception, attention, and working memory stages shown in that model. Long-term memory will also play a role in influencing familiarity with the data being depicted or the underlying graphical form. These are essentially the same as the bottom-up and top-down influences on visual information sampling described in the SEEV model in Chapter 3. Salience and effort are primarily influenced by perceptual, attentional, and working memory stages; expectancy and value are influenced by working memory and long-term memory processes. In general, we will see that less effective task-graph combinations require a longer sequence of mental operations rather than having key task variables represented using easily perceived geometric characteristics.

1.1 Graph Guidelines

We provide five general guidelines for the *construction* of graphs here. We discuss evidence for each guideline in turn. Further guidelines can be found in Gillan, Wickens, Hollands, and Carswell (1998).

1. *Consider the task.* The relative effectiveness of various graph types depends on the task. The graph designer should choose a graphical form that corresponds to task demands.
2. *Minimize the number of mental operations.* The graph designer should try to reduce the number of operations required by choosing an appropriate graph type (e.g., bar graph, pie chart) and arranging information within the graph appropriately.
3. *Use physical dimensions judged without bias.* Perceptual illusions, biases in the judgements of some perceptual continua, and misjudgments of depth can produce error in judgment.
4. *Keep the data-ink ratio high.* Keep the amount of ink that does not depict actual data to a low level.
5. *Code multiple graphs consistently.* Graphs within a set should be designed in a consistent manner.

1.2 Task Dependency and the Proximity Compatibility Principle

There are a large number of tasks people perform with graphs. A convenient taxonomy is shown at the bottom of Figure 4.1 (Carswell, 1992a). In *point reading*, the observer estimates the value of a single graph element. For a *local comparison* the observer compares two values directly shown in the graph. For a *global comparison*, the observer compares quantities that must be derived from other quantities shown in the graph. Finally, for a *synthesis judgment*, the observer needs to consider all data points and make a general, integrative judgment.

In Chapter 3 we introduced the notion of compatibility between the arrangement of multiple information sources on a display, and the task requirements. We saw that this display-cognitive compatibility could be defined in part by the **proximity compatibility principle** (PCP; Wickens & Carswell, 1995). Tasks requiring integration of information are better served by more integral, objectlike displays. The PCP also applies to graphs, as revealed by a **meta-analysis** conducted by Carswell (1992a). The meta-analysis integrated the results of studies in which different graphic formats were compared. Integrated graph types (e.g., a line graph) were compared with more separable formats (e.g., a bar graph or pie chart), as shown in Figure 4.1. Each study was classified by its task demands into one of the four task categories described above, defining a continuum of task proximity. The continuum thus represented the extent to which the integration of all variables was necessary to carry out the task. (See Chapter 3, Section 3.5 and Figure 3.9). Figure 4.2 shows the proportion of studies in each category that showed better performance with the integrated graphs (relative to separated formats), and those that showed the reverse effect. The Figure shows the increasing benefit of integrated graphs as the task required more integration. The comparison of relative effectiveness of tables and graphs (described above) can also be viewed in this manner—a table is highly effective for point reading (focused attention), but less effective for integrative judgments, relative to graphs (Speier, 2006; Vessey, 1991).

As a specific example of the proximity compatibility principle, using the graphs in Figure 4.1, consider this question: How is the rate of growth different between the two factories? Each object (line) of the line graph offers an *emergent feature*—its slope—which can be directly perceived and directly maps to the task (trend estimation). In contrast, the series of pie charts depicts the same data, but no single object represents the rate of growth. The rate must be inferred by comparisons of individual slices over the several years. However, judgments of specific proportion values can be made as well or better with the pie chart than with the line graph.

The PCP also applies to the question of how to label data in a graph. Examine the line graph in Figure 4.1 and ask yourself whether Factory A's production increased from 2010 to 2011. To perform this task, you must first identify the line that represents Factory A. This is not too difficult because labels have been placed in close proximity to the lines. In contrast, if you look at

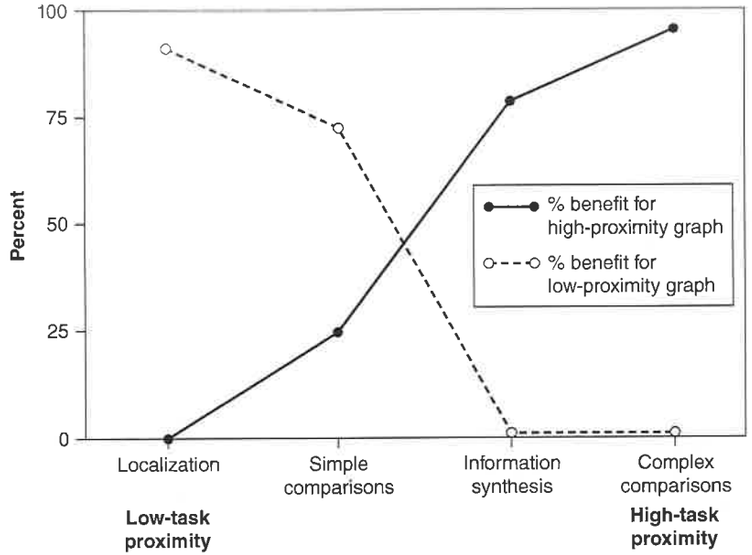

FIGURE 4.2 Proportion of studies showing an object-display advantage (solid line) or disadvantage (dashed line) as a function of task type (focused, left; integrated, right). The figure illustrates the proximity compatibility principle. *Source:* History and applications of perceptual integrality theory and the proximity compatibility hypothesis. University of Illinois Technical Report ARL88-2/AHEL-88-1 Technical Memorandum 8-88.

the bar graph or the pie charts, you need to look for a legend, determine which shading level is assigned to which factory, and remember the coding when you examine the graph again. Several additional mental operations are needed. Thus, a general recommendation is that labels should be placed close to their referents (Gillan et al., 1998).

When a graph shows many variables, direct labels are less feasible. In this case, it is helpful if the order of variables in the legend (going from top to bottom) corresponds to the order in the graph: that is, that the graph and legend are spatially compatible. Huestegge and Philipp (2011) have examined the effect of such compatibility in an experiment in which the eye movements of their participants were measured. Participants were shown a declarative statement (e.g., "In general, people spend more time in front of the computer than the TV") followed by the graph, and their task was to decide if the data shown in the graph were consistent with the statement. They found that when the graph and legend were spatially compatible, less time was required to make the decision.

1.3 Minimize the Number of Mental Operations: Search, Encode, and Compare

When a graph reader examines a graph to accomplish a task, a sequence of perceptual or cognitive operations is performed. Various graphical perception models postulate a general process of *search* (drawing upon attentional processes of visual search as described in Chapter 3), followed by the *encoding* of variables, and ultimately *comparison* of perceived elements with values stored in working memory (e.g., Casner, 1991; Gillan, 1995, 2009; Gillan & Lewis, 1994; Hollands & Spence, 1992, 1998, 2001; Lohse, 1993; Peebles & Cheng, 2003; Pinker, 1990). Each operation is assumed to take time, and have some probability of error. More operations will take more time and will increase the likelihood of error in graph interpretation.

Consider a simple example. Hollands and Spence (1998) found that increasing the number of slices depicted within a pie chart had no effect on response times for judging proportion, whereas increasing the number of bars shown in a bar graph did. The graph reader needs to estimate the whole with the bar graph because no single object represents it. Determining this estimate requires mentally summing the bars: the more bars, the more summation operations, the more time required to perform the task. (Error also increased with more bars.) In contrast, with the pie chart, the entire pie represents the whole and so there is no need for summation operations.

By conducting many studies of this type with particular task-graph combinations, researchers have worked towards general graphical perception models. For example, Gillan (2009) has proposed particular sets of arithmetic and perceptual operations (or **mental operations**) when tasks require simple comparisons, or estimates of differences, sums, ratios, or means, using bar graphs, line graphs, pie charts, and star (object) charts. Gillan summarizes the results of a large number of empirical tests of the model's predictions. Once validated, these general models can then be used to make specific predictions about the time required (or likelihood of error) for a specific judgment.

Visual scanning behavior provides a good measure of the sequences of mental operations. Computational models of graph reading have been developed based on sequences of mental or visual scanning operations (e.g., Chandrasekaran & Lele, 2010; Peebles & Cheng, 2003). The formal aspect of the models also helps in comparing human performance to some optimal level. For example, Peebles and Cheng found that their participants unnecessarily revisited certain graph locations as they executed the task, demonstrating non-optimal scanning patterns. It would appear that as the users scanned the graph they forgot information accessed from the graph earlier (a failure to adequately encode a value). A redesigned graph might avoid this problem.

In many everyday situations, the graph reader's task might simply be to ask, "What is this graph saying?"; that is, to synthesize the graph's message as a whole. Such integration tasks have been shown to be carried out in steps (Carpenter & Shah, 1998; Ratwani, Trafton, & Boehm-Davis, 2008). In particular, eye movement and verbal protocols indicate that people segregate the graph into chunks or visual clusters (e.g., light and dark bars in the bar graph in Figure 4.1). Eye movements are often focused on the boundaries between the clusters, segregating the graph into different parts, which can then be compared. The result of the comparison often lead to a cognitive integration of the graph's message (e.g., Factory B's production advantage keeps getting bigger). Such processing can be aided by ensuring that visual clusters are easily distinguishable (e.g., by using color coding or shading, discussed in Section 2), but the graph should not encourage the formation of too many visual clusters by having too many uniquely coded variables (Ratwani et al., 2008).

In summary, the graph designer should always strive to reduce the number of operations by first choosing an appropriate graph type and then arranging information within the graph appropriately. In PCP terms, reducing the number of operations reduces information access cost. Various models instruct how this should be done.

1.4 Biases in Graph Reading

In particular situations, the judgments people make in extracting information from graphs are biased (Gillan et al., 1998). That is, people systematically overestimate (or underestimate) quantities relative to their true values. Some biases are related to optical illusions that distort our sense of perception. For example, when viewing the **Poggendorf illusion**, shown in Figure 4.3a, people "flatten" the sloping lines horizontally. The same illusion tends to flatten the slope of a line in a line graph, as indicated by the arrows in Figure 4.3b. Thus, a point far from the axis (e.g., a point

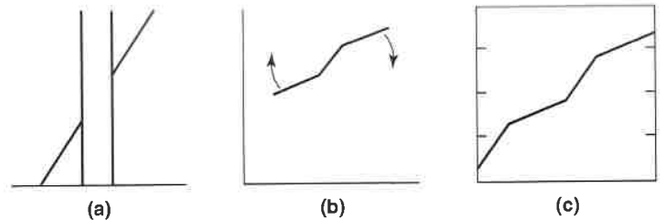

FIGURE 4.3 (a) The Poggendorf illusion: the two diagonal lines actually connect. (b) A line graph susceptible to "bending" from the Poggendorf illusion. (c) Debiasing of the Poggendorf illusion by marked edges on both sides. *Source:* E. C Poulton, "Geometric Illusions in Reading Graphs," *Perception & Psychophysics, 37* (1985), 543. Reprinted with permission of Psychonomic Society, Inc.

on the right side of the line shown in the figure) will tend to be underestimated (Figure 4.3b; Poulton, 1985). Poulton found that the illusion is greatly reduced if a graduated axis is provided on each side (Figure 4.3c). Gridlines placed within the graph are also helpful in reducing the bias (Amer, 2005).

A second example of bias occurs when comparing *differences* between two lines of different slope (Cleveland & McGill, 1984). The vertical difference between the two curves in Figure 4.4 is actually smaller on the left. Yet perceptually the difference appears smaller on the right because judgments of differences along the y-axis are biased by the visual separation (or Euclidean distance) between the two curves rather than the vertical separation. One solution is to plot the differences directly (Figure 4.4, bottom).

Other biases result from perceptual limitations in judging areas and volumes, which are commonly used to represent quantity in graphs. Volume is becoming especially prevalent given the frequent use of 3D graphical formats (Carswell, Frankenberger, & Bernhard, 1991; Siegrist,

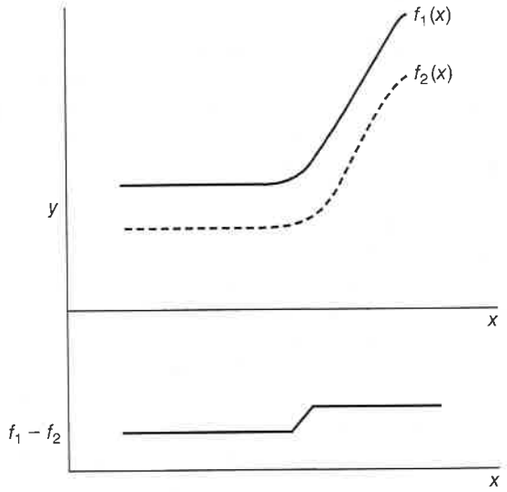

FIGURE 4.4 Biases in perceiving differences between pairs of lines $f_1(x)$ and $f_2(x)$ with changing slopes. The bottom curve plots the difference $f_1(x) - f_2(x)$, which is larger on the right than on the left.

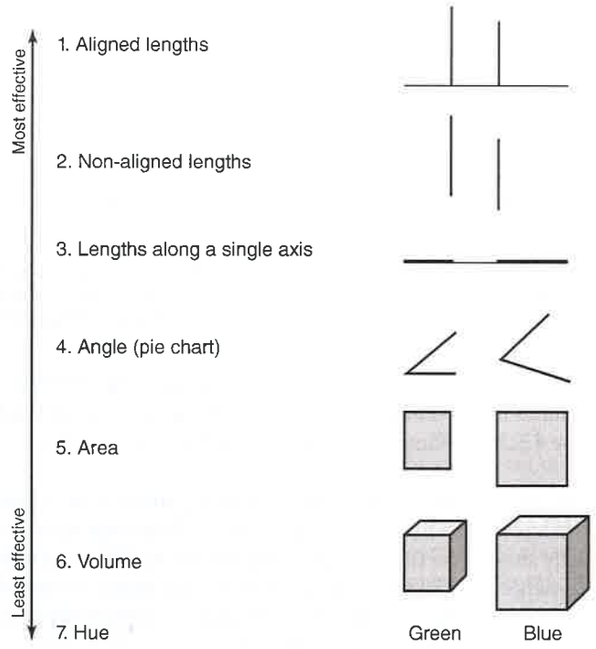

FIGURE 4.5 Seven graphical methods for presenting quantities to be compared. The graphs are arrayed from most (top) to least effective (bottom).

1996; Spence, 2004). Based on a set of experiments they conducted, Cleveland and McGill (1984, 1985, 1986) proposed that our ability to make comparative judgments of two quantities in a graph progressively degrades in the order shown in Figure 4.5. The best comparative judgments are made with the evaluation of two linear scales, aligned to the same baseline. (We made a similar point in Chapter 3, when we considered how aligning bars to the same baseline created the emergent feature of slope.) The poorest judgments occur when people compare two areas, volumes, or color patches. The Cleveland and McGill ranking shown in Figure 4.5 provides a useful framework for a graph designer and corresponds to the predictions of the PCP for focused tasks like local comparison and point reading (Carswell, 1992b).

The ranking in Figure 4.5 is likely related to observed biases in judging **perceptual continua** (types of stimuli). When people estimate magnitudes by assigning numbers to various sizes of objects (the **magnitude estimation** procedure developed by Stevens, 1957), they show certain biases. Some continua, like area and volume, produce **response compression**: each unit increase in physical magnitude causes less and less increase in perceived magnitude. Other stimuli, such as color saturation, tend to show **response expansion**: each increase in physical magnitude causes incrementally greater increases in perceived magnitude. Lengths tend to be judged with little bias. Stevens (1957, 1975) found that the relation between physical and perceived magnitude can be expressed by the power function called **Stevens' law**, with the exponent representing the amount of response compression or expansion. When the exponent is less than 1.0, response compression occurs; when it is greater than one, response expansion occurs; when it is equal to 1.0, no bias occurs. Estimates of the areas and volumes shown in graphs are thus subject to response compression so

that large areas and particularly volumes will tend to be underestimated. In general, the use of areas, volumes, color saturation, and other perceptual continua whose Stevens' exponent differ from unity should be avoided in graphs.

Moreover, the bias described by Stevens' law affects more complex judgments where multiple quantities are involved, such as judgments of proportion (e.g., what proportion is A of B?; Hollands & Dyre, 2000). Suppose you were asked to divide a horizontal line into two parts corresponding to two slices of a pie, as shown in Figure 4.6a. When judging graphs (e.g., pie charts, stacked bar graphs) depicting proportion, people tend to show cyclical bias patterns (e.g., overestimation from 0–.25, underestimation from .25–.75, overestimation from .50–.75, and underestimation from .75 to 1). The "amplitude" of the cyclical pattern is determined by the Stevens' exponent (for the pie charts shown in the figure, the estimated exponent was less than 1.0), and the "frequency" of the bias pattern was determined by the number of available tick marks (compare the upper panels of Figure 4.6). When tick marks are added to the graph, as shown in Figure 4.6b, the bias frequency doubles, reducing error. Intermediate reference points (the tick marks) are used by observers to subdivide the graph into components, which has the beneficial side effect that error is reduced, even as the Stevens exponent stays constant.

In summary, bias in making relative judgments with graphs can be reduced by: 1) avoiding continua whose Stevens exponents differ from 1.0; and/or 2) making reference points available (e.g., adding tick marks). It is possible to make less effective perceptual continua (e.g., area) more effective by adding reference points to the graph.

1.5 The Data-Ink Ratio

As we noted earlier, graph readers naturally scan or search through the available graphical elements. This is especially true in situations where the reader is unfamiliar with the graph type or is otherwise inexperienced (Peebles & Cheng, 2003). In Chapter 3 we learned that unnecessary

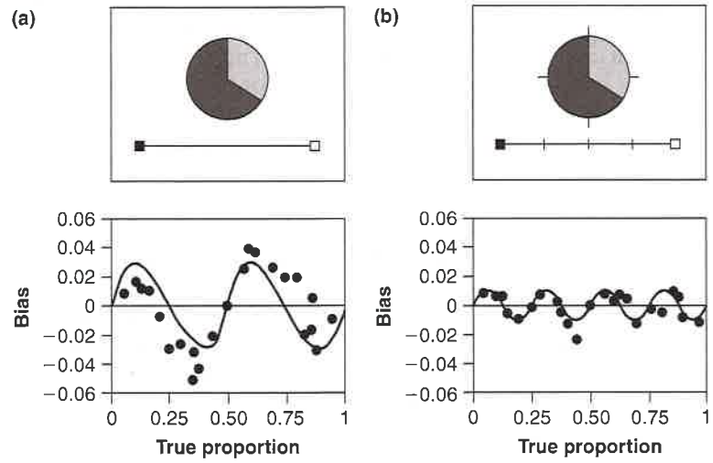

FIGURE 4.6 Patterns of cyclical bias in judging graphs. (a) Bias as a function of true proportion for pie charts. (b) Bias as a function of true proportion when tick marks are added. The bias pattern changes from two to four cycles and overall error is reduced. The curved functions show the predictions of the cyclical power model (Hollands & Dyre, 2000), derived from Stevens' law.

visual elements (clutter) will slow visual search. The greater the number of visual elements in the graph, the greater the number of scans required. Graph designers should therefore strive to eliminate those extraneous elements of the graph that do not carry information (Wang, 2011). In an influential book, Tufte (2001) distinguished between the ink in a graph used to portray data and superfluous non-data ink. He argued for a **data-ink ratio** principle, which states that the amount of ink that does not depict data points should be kept to a minimum (Tufte, 2001).

In line with the principle, techniques have been developed to modify graphical elements so that more data can be portrayed in the same amount of space, without sacrificing judgment accuracy (Heer, Kong, & Agrawala, 2009). The higher the data-ink ratio (i.e., more ink associated with data and less unnecessary ink), the faster the time to make a variety of judgments, and the greater the accuracy (Gillan & Richman, 1994). In addition, integration tasks (e.g., global comparison, synthesis judgments) appear to be more affected than focused tasks by the data-ink ratio. Gillan and Richman's results also suggest that the use of pictorial backgrounds (e.g., the picture of a bank behind a bar graph depicting financial data, in the typical *USA Today*-style graph) is particularly damaging, especially for more integrated judgments. Similarly, Renshaw, Finlay, et al. (2004) compared a 2D line graph with a 3D ribbon graph (lines were represented as ribbons viewed from an oblique angle), and found performance advantages for the 2D format, which had a much higher data-ink ratio. Ratwani et al. (2008) found that task-irrelevant labels required extra fixations and increased comprehension time; when the labels were removed the extra fixations and time penalty were eliminated. Thus, there is good evidence to suggest that the use of high data-ink ratios will, by reducing distraction (failure of focused attention), make a graph more effective, especially for integration tasks, and that non-data ink should be eliminated from graphs. This is especially important to remember given that people appear to prefer graphs with more non-data ink (Inbar, Tractinsky, & Meyer, 2007).

It is possible to carry the data-ink ratio principle too far, however (Carswell, 1992b; Wickens, Lee, Liu, & Gordon-Becker, 2004). The lines connecting points within a line graph represent non-data ink (data are fully represented by the points). But deletion of the lines is not always a good idea because, as we saw in Chapter 3 and also in Figures 4.1 and 4.2, the line slope serves as an *emergent feature*. Limited use of non-data ink can be useful in helping the user interpret graphical elements (Gillan & Sorensen, 2009). If the non-data imagery is linked to the content of the graph, it can be effective in making the graph more distinctive, and therefore more memorable (Bateman et al., 2010). In general, then, non-data ink should be avoided, but if used judiciously some non-data ink may assist in graph comprehension.

1.6 Multiple Graphs

The previous discussion has focused on the ideal, compatible properties of single graphs. An equally important issue lies in the presentation of linked or multiple graphs, which may show related sets of data (e.g., one graph shows the prevalence of several diseases for men, the other for women). This is analogous to the interactive display or information visualization situation where the data are complex enough to require viewing in multiple formats or windows to understand their interrelation (Chen et al., 2007). Here the graph designer should consider the relationship between successively viewed graphic formats, in addition to the optimization of each format by itself. Four specific concerns can be identified.

1. *Coding Variables.* Shah and Carpenter (1995) have shown that our mental representation of coded variables (different lines) is qualitative or nominal, whereas our representation of variables placed along the *x*-axis of the graphs is in quantitative metric terms. This has two implications for multiple graph construction: 1) Build the graphs so that quantitative variables are placed on the *x*-axis; and 2) If all variables are qualitative, build the graphs so that the most important differences are encoded as the variables represented by the two (or more) points on each line along the *x*-axis, since we seem to be most sensitive to these changes. In this way, the variable's effect is directly represented by an emergent feature—the slope—of the constructed graphs. The *differences* in slope (the angle between two lines) serves as an emergent feature, as noted in Chapter 3.
2. *Consistency.* When the same data are plotted in different ways, it is important to maintain consistency across graphs (Gillan et al., 1998). For example, the variable coded by line type (e.g., dashed versus dot) in one graph, should, where possible, be coded by the same physical distinction in all graphs. If such consistency is needlessly violated, the reader will need to exert greater cognitive effort (i.e., a longer sequence of mental operations) to switch from one graph to the other. High consistency creates good **visual momentum** as the eye moves from graph to graph (Woods, 1984), a concept considered in the next chapter.
3. *Highlighting differences.* When related material is presented, it becomes critical to highlight the *changes* from graph to graph, either prominently in the legend or in the symbols themselves. For example, a series of graphs presenting different Y variables as a function of the same X variable should highlight the Y label. This system allows the same cognitive set to be transferred from graph to graph, while the single mental revision that is necessary is prominently displayed. The time- and effort-consuming visual search necessary to locate the changed element is minimized (Gillan et al., 1998), reducing information access cost.
4. *Short distinct legends.* Legends of similar graphs should highlight the distinct features, not bury them as a single word that is nearly hidden in the middle or end of otherwise identical multiline legends. Unfortunately, word processors make it all too easy to copy a long legend from one graph to another, making it difficult to detect each graph's unique features. The caption for each graph should be written in short, efficient language that highlights the differences among graphs.

In conclusion, even though graphs are relatively simple, static displays, meant to be interpretable by the layperson, there are a number of significant design issues to consider. We shall see many parallels when we consider interactive information displays in the next section and next chapter, because the same digital or analog representations are often used. The analog representations often take similar forms, with geometric and spatial elements being used to represent the value of variables of interest in a similar manner. Thus, overarching principles (such as the proximity compatibility principle and consistency) will reappear as we consider such displays. However, with information displays the situation is dynamic, the data are real-time or close to it, and the operator is often in the position of controlling some of the variables being portrayed in the display (or overseeing automation that is controlling the variables). This was the *supervisory control* task described in Chapter 3. The control of such variables often requires significant training or experience (e.g., controlling a nuclear power plant, flying an aircraft). In contrast, graphs are usually designed to be interpretable by the layperson. Thus, in terms of our information processing model (Figure 1.1), the use of feedback from the environment (after control actions) takes on an important role. Displays need to represent the right variables in an intuitive manner to provide a useful guide for action (Bennett & Flach, 2011). We consider these topics in the next section.

2. DIALS, METERS, AND INDICATORS: DISPLAY COMPATIBILITY

Many dynamic systems controlled by human operators present information in dynamic analog form, using dials, meters, or other changing elements, to represent the momentary state of some part of the system. It is important that dials and meters be compatible with the operator's **mental model** of the system. The mental model, a concept we will discuss further in Chapter 7, forms the basis for understanding the system, predicting its future behavior, and controlling its actions (Gentner & Stevens, 1983; Moray, 1998; Park & Gittelman, 1995; St-Cyr & Burns, 2001). As a consequence, there are three levels of representation that must be considered in designing display interfaces, as shown in Figure 4.7: (1) the physical system itself; (2) the user's mental model; and (3) the interface between these two, the display surface on which changes in the system are presented to the operator, and which help form the basis for control action and decision (Bennett & Flach, 2011). It is important to maintain a high degree of *compatibility* among all three representations.

In achieving this compatibility, it is first important that the properties of the interface accurately reflect the dynamics of the physical system, a correspondence referred to as **ecological compatibility** (Vicente, 1990, 1997). This will help the operator's mental model to correspond better to the physical system dynamics (St-Cyr & Burns, 2001; Vicente, 1997). Such correspondence will be aided by displays that show the key physical parameters in effective and intuitive ways, as well by good operator training, discussed in Chapter 7. Second, **display compatibility** is achieved by display representations whose structure and organization are compatible with the user's mental model.

Given the increase in the use of automation in complex systems (discussed in Chapter 12), the physical representation includes not only the system performing the physical work, but also any automated system controlling the process. Thus, for example, the physical system for an aircraft includes not only the rudder, engines, elevators, and ailerons but also the automated systems used to control those aircraft components. It is important for the mental model to reflect the automated systems correctly in order to maintain appropriate awareness should the system fail. For example, Sarter (2008) noted that gaps and misconceptions in pilots' mental models of flight

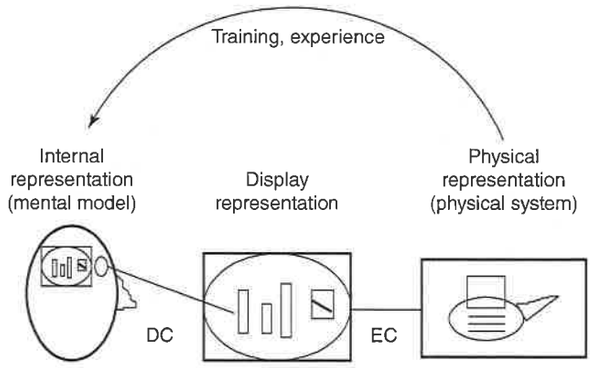

FIGURE 4.7 Representations of a physical system. Two types of compatibility are portrayed: that between the physical system and a display (ecological compatibility: EC) and that between the display and the user's mental model (display compatibility: DC). The Figure also highlights the importance of training to the influence of the physical representation upon the mental representation.

deck automation in the Boeing B737 and Airbus A320 contributed to errors made by those pilots. Recent aviation accidents like Colgan Air Flight 3407 near Buffalo, New York (Sorensen, 2011) were at least partially attributable to the pilots' lack of understanding of what the automation was doing when the plane lost control.

When considering display compatibility, it is important to distinguish between *analog* or *continuous* systems and *digital* or *discrete* systems. In general, analog systems are those whose behavior is governed by the laws of physics, and therefore change continuously over time (e.g., controlling an aircraft, a ground vehicle, or an energy conversion process). The physics defines an *ecology*, and hence makes ecological compatibility important. In considering analog systems, it is important to distinguish between *static* and *dynamic* components of display compatibility. We now consider each in turn.

2.1 The Static Component: Pictorial Realism

The **principle of pictorial realism** (PPR; Roscoe, 1968) has two parts. The first part can be defined as follows: if a variable's physical representation is analog, then its display representation should also be analog (Roscoe, 1968). The representation of aircraft altitude is a typical instance. Physically, altitude is an analog quantity, with large changes in altitude more important than small changes. Conceptually, the pilot likely represents altitude in analog form. Therefore, to achieve compatibility, a display of altitude (i.e., an altimeter) should be in analog format (e.g., a needle position changing on the display to indicate a change in altitude) rather than digital. The human transformation of symbolic digital information to analog conceptual representation imposes an extra cognitive processing step, leading to longer visual fixations, increasing processing time, and increasing the likelihood of error (Grether, 1949).

There are, of course, other factors that influence the choice of analog or digital representations of altitude or of other continuously varying quantities. The nature of the user's behavioral response—which is often driven by task requirements—matters. Miller and Penningroth (1997) had participants read analog and digital clocks and report the time in different ways. When they were asked to read the time as exact numbers (e.g., 2:40→"two forty") the digital format was found to be superior. On the other hand, the need to estimate at a glance the distance of that variable from some limit by stating minutes before the hour (e.g., 2:40→twenty minutes to three) favored the analog format. Similarly, perceiving the magnitude of a variable when it is rapidly changing or determining rate-of-change or event onset information favors an analog representation (Proctor & Van Zandt, 2008; Schwartz & Howell, 1985). Given the flexibility of electronic displays, it is common to use both formats within a single display. This meets the needs of multiple tasks. For example, in general an analog representation is effective for representing heading to a soldier using a head-mounted display while wayfinding in an unfamiliar environment (Kumagai & Massel, 2005). This follows the principle of pictorial realism. Nonetheless, it is useful for the display to show additionally the specific heading to a waypoint digitally, to aid the soldier who is verbally communicating a heading to another soldier.

There are many variables whose internal representations are likely analog (e.g., temperature, pressure, speed, power, or direction). In addition, some conceptual dimensions have the characteristic of an ordered quantity with multiple levels (e.g., degree of danger or readiness status); these will also likely benefit from analog representation.

The second part of the PPR is that the *direction* and *shape* of the display representation should be compatible with the mental (and physical) representations. Consider a violation in *direction*: an altimeter that places high altitudes low on the display, and vice versa. While this would still be an analog representation, our mental model of altitude mimics the physical variable

itself: high altitudes are up and low altitudes are down. Therefore, the altimeter should present high altitudes at the top of the scale and low ones at the bottom. Analogously, high temperatures should be placed higher, and low temperatures lower, on a display.

Display compatibility may be violated in terms of *shape* if a circular altimeter (pointer or dial) represents the vertical and linear conception of altitude (Grether, 1949). The PPR is also violated by dissecting a single, continuous variable into separate parts. Grether reported that operators had a more difficult time extracting altitude information from three concentric pointers (indicating units of 100, 1,000, and 10,000 feet) than from a single pointer. In sum, displayed quantities should correspond to the operator's mental model of them, which in turn reflects characteristics of the physical world. The concept of static compatibility may also be applied to systems that are not inherently analog, but have some ordered spatial component, such as an expert system's decision logic, or a circuit diagram.

When we talk of pictorial realism in the PPR, it is important to understand that we are not arguing for blind acceptance of realism in displays; that is, to assume that realism is always a good thing. Smallman and St John (2005; Hegarty, Smallman, & Stull, 2012) labeled this misplaced faith in realistic information display as **naïve realism**. Smallman and Cook (2010) showed users photorealistic three-dimensional terrain models, as well as less realistic topographic maps of the same terrain. Their participants rated the models as more realistic than the topographic maps, and also thought that they would perform better with the more realistic displays. However, the participants actually performed *worse* with the more realistic terrain models because the greater realism meant that extraneous data were shown along with task-relevant information. The user is faced with the burden of additional cognitive effort to extract the task-relevant information from the extraneous data (or alternatively, filter out the non-relevant data). In contrast, here we have been arguing that the display representation should be compatible with the user's mental model as she performs a task. Any particular task in an analog system will demand that certain parameters are attended to while others are not relevant. The PPR argues for analog representation of these key parameters, whereas naïve realism would argue that all domain parameters be explicitly represented, even those that are not relevant to the current task activity.

2.2 Color Coding

Before turning to a discussion of dynamic aspects of display compatibility, it is important to consider another static form of display compatibility: the role of **color** in display design. We discussed color coding in Chapter 2, in terms of absolute judgment, and in Chapter 3 in terms of its attentional impact in visual search and the proximity compatibility principle, and we will reconsider color coding when we discuss its role in information visualization (Chapter 5). We summarize here several characteristics of color that have practical implications for display design.

- A unique color stands out from a monochrome background, and as we saw in visual search, also allows for rapid parallel search for a target (Christ, 1975).
- **Color hue** is useful for coding *categorical* or *qualitative* information (e.g., blue and red symbols on a map to show friendly and hostile forces). However, like other sensory continua, color is subject to the limits of absolute judgment (see Chapter 2). Thus, the system designer should probably use no more than about seven hues in a display (Carter & Cahill, 1979; Flavell & Heath, 1992). In conditions where ambient light varies (e.g., in a cockpit or hand-held display), absolute judgment performance will likely be impaired (Stokes et al., 1990) and fewer then seven levels are strongly recommended.

- Color hue is effective for segregating categories of objects within a display (Yamani & McCarley, 2010), and for showing discrete state changes (Smith & Thomas, 1964; Van Laar & Deshe, 2007).
- Certain colors have well-established symbolic meaning within a population (e.g., red is often used to indicate danger, or stop; green signals safety, or go). Because these sometimes vary across culture (Courtney, 1986), such coding is often referred to as a **population stereotype**, discussed further in Chapter 9. Coding levels should not conflict with population stereotypes (e.g., assigning red to "go" or "safe").
- Color hue does *not* generate a natural ordering (i.e., from "most" to "least" in a way that lends itself to analog displays (Merwin, Vincow, & Wickens, 1994). Red is not perceived as "more" or "less" than green. Thus, color hue is not effective for **relative judgment** or **comparison** tasks in which users are comparing values along a continuous or ordinal scale, such as deciding which values is greater or less, which is of course important for the representation of analog variables. **Color saturation** is more effective for this purpose (Bertin, 1983; Kaufmann & Glavin, 1990). Ordered **brightness** scales have also been shown to be more effective than scales based on hue variation (Breslow, Trafton, & Ratwani, 2009; Spence & Efendov, 2001; Spence, Kutlesa, & Rose, 1999) for relative judgment tasks.

There is evidence to suggest that judicious combinations of hue and brightness can be effective for both identification and comparison tasks. For example, Spence et al. (1999) showed that ordered color scales in which brightness was covaried with hue produced more accurate comparison judgments than brightness variation alone. An algorithm called Motley has been developed to produce color scales varying in both hue and brightness (Breslow, Trafton, et al., 2010). These authors showed that both identification and relative comparison tasks were well served by Motley's ordering. Thus, by clever combination and selection of display elements, it is possible to design a display that serves multiple purposes well. We shall return to this **hybrid display** concept when we discuss display movement in the next section.

2.3 Compatibility of Display Movement

If motion is occurring in the physical system itself, it can be useful to represent that motion by display motion (rather than by using static displays) to produce an appropriate mental model of the situation (Park & Gittelman, 1995). Beyond that, however, the compatibility of *direction* between the display and the mental model is also important. Roscoe (1968) and Roscoe, Corl, and Jensen (1981) proposed the **principle of the moving part** (PMP)—that the direction of movement of an indicator on a display is compatible with the direction of movement of an operator's mental model of the variable. In the case of the mercury thermometer, this principle is typically adhered to because a rise in the height of the mercury column indicates a rise in temperature. There are, however, circumstances in which the PMP and the PPR operate in opposition, and so one or the other must be violated.

An example of this violation is shown in Figure 4.8, which could represent an altimeter. In the **moving-pointer display** (Figure 4.8a) both principles—moving part and pictorial realism—are satisfied. High altitude is at the top and an increase in altitude is indicated by an upward movement of the moving element on the display. However, this simple arrangement can only show a small range of altitudes or requires an extremely compressed scale where motion would be barely visible. One solution is to have a fixed pointer and move the display scale when necessary to show only the relevant part (a **moving-scale display**; Figures 4.8b and c). If the moving scale is designed to follow the PPR, high altitudes should be at the top of the display (Figure 4.8b).

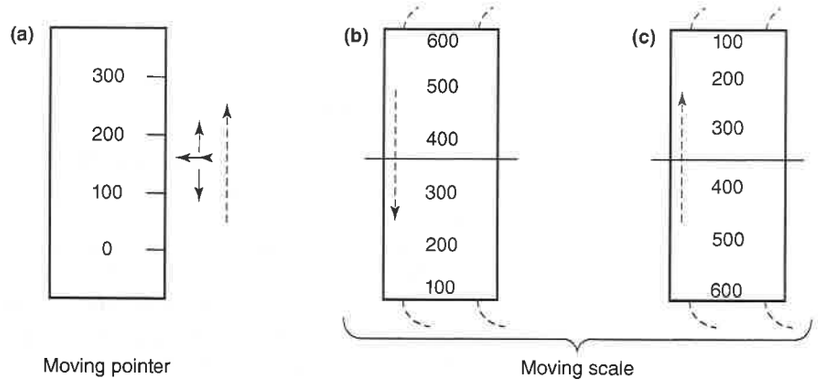

FIGURE 4.8 Display movement. (a) Moving-pointer altimeter; (b) and (c) are moving-scale or fixed-pointer altimeters. The dashed arrows show the direction of display movement to indicate an increase in altitude.

However, this means that the scale must move *downward* to indicate an *increase* in altitude—a violation of the PMP. If the labeling is reversed to conform to the PMP (Figure 4.8c) this change will reverse the orientation and display high altitude at the bottom, violating the PPR! A disadvantage for both moving-scale displays is that scale values become difficult to read when the variable is changing rapidly since the digits themselves are moving.

A possible solution here is to employ a hybrid display. The pointer moves as in Figure 4.8a, but only a restricted portion of the scale is exposed. When the pointer approaches the top or bottom of the window, the scale shifts more slowly in the opposite direction to bring the pointer back toward the center of the window, and expose the newer, more relevant region of the scale. Thus the pointer moves at higher frequencies in response to the more salient aircraft motion and the scale shifts at lower frequencies as needed. This way both principles—pictorial realism and moving part—are satisfied. Head-up displays (described in Chapter 3) often use this approach to show altitude.

Or consider the traditional aircraft attitude indicator (or artificial horizon display), which shows the aircraft's orientation in space (an aircraft's attitude includes roll, pitch, and yaw, but here we will concentrate on roll, when the wings dip left or right). Here, a stable aircraft is positioned relative to a moving horizon (see Figure 4.9a). This looks like what the pilot sees through the aircraft window (because of this, it is sometimes referred to as an **inside-out display**), and therefore conforms to the PPR. But when the plane rotates (rolls or banks) it is the horizon not the aircraft that moves. This violates the PMP because pilots perceive the world as stable and the aircraft moving through it (Johnson & Roscoe, 1972). Furthermore, the horizon will rotate in an opposite direction to the aircraft, hence inviting confusion and an incompatible response (Roscoe, 2004). As above, constructing the display so that the aircraft moves and the horizon is stationary (an **outside-in display**) produces the opposite problem. It violates the PPR, since the static picture that is drawn (horizontal horizon, tilted airplane) is incompatible with what the pilot perceives through the window (tilted horizon, horizontal airplane).

A hybrid display called the **frequency separated display** (Figure 4.9(c); Lintern, Roscoe, & Sivier, 1990), like the hybrid altitude scale above, captures the best of both worlds, conforming in different ways to both principles. Rapid movement of the aileron (controlling roll or bank) will cause the aircraft symbol to roll in the same direction of the control, conforming to PMP. However,

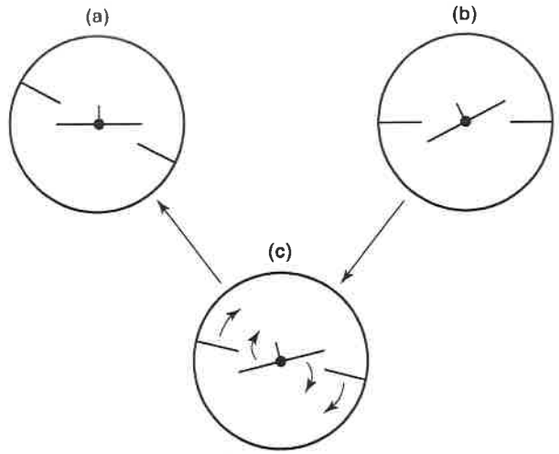

FIGURE 4.9 Aircraft attitude display. (a) inside-out, (b) outside-in, and (c) frequency-separated display. All displays show an aircraft banking left. Low-frequency return to steady state is indicated by arrows in (c).

following a relatively sustained roll, or slow roll back to level, the horizon rotates to the new orientation, as the plane symbol rotates with it, back to horizontal, hence restoring the correct "picture" of what the pilot sees when looking forward: conforming to the PPR. Thus the rapid motion conforms to the PMP, while the slower "steady state" conforms to the PPR. Evaluations with skilled pilots have shown the success of frequency separation over displays that follow a single principle (Beringer, Williges, & Roscoe, 1975; Ince, Williges, & Roscoe, 1975; Roscoe & Williges, 1975). Thus, the frequency-separated display illustrates a more general principle: sometimes clever design can produce a system that adheres to two apparently contradictory principles with effective results.

Another type of frequency-separated display is called a **tethered display** (Wickens & Prevett, 1995). Consider a gaming environment in which a user controls a virtual avatar in a three-dimensional world. It is quite common in such environments to have the viewpoint placed behind and above the avatar, and connected to it, so that when the avatar moves the viewpoint moves with it in "tethered" fashion. The use of similar technologies is being explored for remote vehicle control (Hollands & Lamb, 2011; Wang & Milgram, 2009). Wang and Milgram developed a virtual tether with dynamic properties so that there is gradual adjustment of the camera's viewing position after a movement by the avatar. Importantly, a dynamic tether can be constructed so that, like the two hybrids described above, the tether acts first as an inside-out display, with the control motion first affecting the avatar motion in the same direction, and then a compensatory motion of the surrounding scene occurs (outside-in display). Dynamic tethers based on this frequency-separated principle were shown by Wang and Milgram to be superior to rigid tethers (which would be likened to an inside-out display) for controlling the motion of a virtual aircraft through a curved tunnel.

2.4 Display Integration and Ecological Interface Design

The PPR suggests that an array of displays should be spatially compatible or congruent with the array of physical components that they represent, as illustrated in Figure 4.7. However, as discussed in Chapter 3, there are other ways of integrating information on displays to be compatible

with the operator's need to mentally integrate that information, such as the proximity compatibility principle (Wickens & Carswell, 1995). We also noted that many creative design solutions can configure display elements to produce emergent features, when those elements change in certain critical ways that are relevant to the operator's task. When this configuration is done in a way to reflect the constraints of the natural physical system being represented, the resulting displays are called **ecological interfaces** (Vicente & Rasmussen, 1992; Vicente, 2002), conforming to *ecological compatibility*. In this section we will focus on such interfaces.

Interfaces based on the principles of ecological interface design have been developed and assessed in a large variety of work domains. These include nuclear process control (Burns et al., 2008, Burns & Hajdukiewicz, 2004), petrochemical systems (Jamieson, 2007), medical anesthesia (Jungk, Thull, Hoeft, & Rau, 2001), semiconductor manufacturing (Upton & Doherty, 2007), military command and control (Bennett, Posey, & Shattuck, 2008), and the separation of aircraft in free flight (Van Dam, Mulder, & van Paassen, 2008). One of the key features of ecological displays is that they are the result of a process in which the work domain is analyzed not just in terms of its physical form (e.g., pipes and valves), but also in terms of function (what is the purpose of the system), and at an abstract level (what is the physics of the system). Key variables that the operator needs to consider become apparent to the human factors designer through this **work domain analysis** (Burns & Hajdukiewicz, 2004; Vicente, 1999).

For example, in the context of nuclear process control, Burns et al. (2008) compared an ecological display to a traditional display. The traditional display showed the equipment (turbines, valves, pipes) with individual process values (pressure readings, valve positions) in numeric form. In contrast, the ecological display mapped important conceptual variables like mass flow balance to emergent features of the display. For example, as shown in Figure 4.10a, two bars were used to represent the masses of two fluids. A line was drawn between the bars, with the center of the line indicated by a hatch mark; a bubble was placed on the line that acted like the bubble on a carpenter's level. If the two masses were equal then the bubble was found at the hatch mark; if the mass on the left was less than the right, the bubble moved to the right away from the hatch mark, and vice versa (Lau et al., 2008). Furthermore, the emergent feature of the line slope represented the mass balance so that when the mass output from one subsystem was equal to the total mass pumped the line was level, but if flow balance for a set of valves was greater or less than a critical value the line sloped to the left or right at an angle proportional to the disparity. Burns et al. showed that these ecological displays were more effective than the traditional displays for detecting unexpected system failures.

As another example, Seppelt and Lee (2007) developed an ecological interface for an adaptive cruise control (ACC) system. These systems adjust the brake or throttle to maintain a constant distance from the driver's vehicle to a vehicle in front. ACC systems have braking and sensor limitations, which means that the driver must intervene (i.e., hit the brakes) in some situations. The display developed by Seppelt and Lee mapped the physical variables that the driver must monitor and control to certain characteristics of the display. The physical variables included the difference between the velocities of the two vehicles, the distance between the vehicles (scaled to the velocity of the driver's vehicle), and the estimated time to collision (which we discuss later in the chapter). The particular mapping they used meant that the shape of the display changed depending on whether the situation was potentially hazardous or not. If the driver's vehicle was approaching the vehicle in front too quickly, a triangular shape (like a yield sign) was produced; if the vehicle in front was traveling more quickly than the driver's vehicle, the display looked like a trapezoid (empty road ahead) instead, as shown in Figure 4.10. Thus, the emergent feature of shape was directly mapped on to the driver's task of working with the automation to ensure an

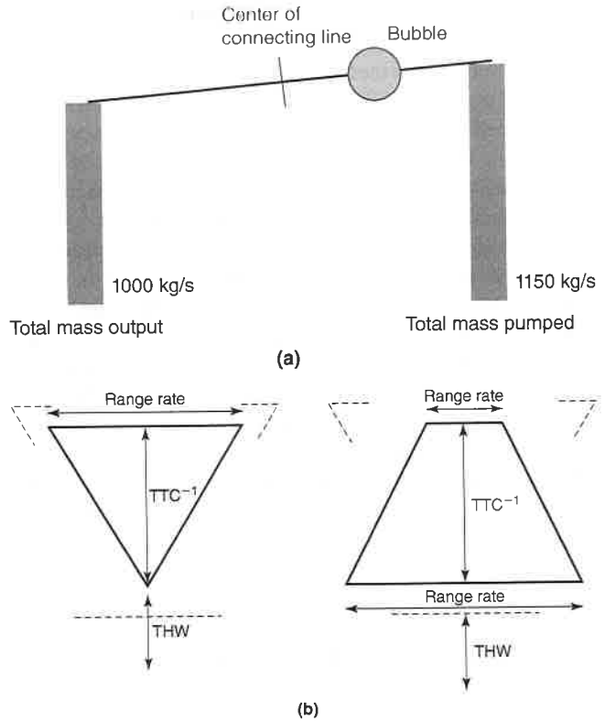

FIGURE 4.10 Examples of ecological displays. (a) Carpenter's level display (Lau et al., 2008). When the two bars are not equal (as they should be), the bubble deviates by shifting away from the hatch mark. (b) Adaptive cruise control display (Seppelt & Lee, 2007). The triangular yield shape on the left indicates that the driver should brake; the trapezoid on the right indicates a safe following distance. TTC = time to collision. THW = time headway (distance from car in front divided by own car velocity).

appropriate following distance. Seppelt and Lee showed that having this ecological display helped drivers maintain the correct following distance (relative to without the display) in situations with both rain and traffic.

There has been considerable effort put into how to generate the most effective displays based on the principles of ecological compatibility. While ecological interface design (EID) provides general guidelines for displays, there are often multiple display options that could meet the guidelines. Indeed, Vicente (2002) has argued that the benefit of EID is not only attributable to the specifics of the functional form, but also that important functional information is available to support the operator's cognitive activities. Jessa and Burns (2007) evaluated particular ecological display options for three different display-reading activities: determining target levels, determining a change in direction, and interpreting proportions. They found that for target value indication, a bull's eye shape (an *object display* in which a solid circle was centered in a larger empty circle) was most effective; for changing direction, a display that showed values on either side of a vertical zero line was most effective, and for depicting the ratios between quantities a bar graph (in which smaller values were shown in proportion to a set of larger values) was most effective.

Jessa and Burns showed that the effectiveness of their ecological displays was determined by the judgment task being performed: integrated tasks (e.g., determining overall status or ratios among various variables) were performed best by displays that integrated those values into a single object, and a focused task (determining if multiple individual variables were greater or less than zero) was best performed a separated format. These results are consistent with the proximity compatibility principle.

Given the importance of proximity compatibility and the form of the task representation (Zhang & Norman, 1994) to display design, we have modified Figure 4.7 to incorporate proximity compatibility as well, as shown in Figure 4.11. The set of compatibility principles shown in Figure 4.11—display, ecological, and proximity compatibility—offer in combination a validated set of display guidelines, one of the most powerful frameworks in the engineering psychology of display design. We will revisit compatibility in the context of information visualization in Chapter 5, display modality in Chapters 6, 7, and 9, and motor responses in Chapter 9. Displays that are compatible in these various respects are read more rapidly and accurately than incompatible ones under normal conditions. More important, their advantages increase under conditions of stress (see Chapter 11). The four representations in Figure 4.11 are tightly intertwined in a successful system; this congruence is most likely to occur when the three types of display compatibility are met.

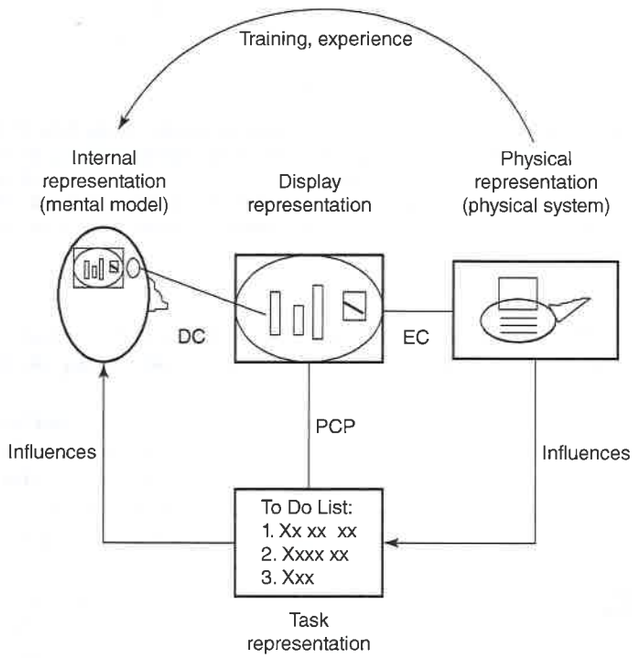

FIGURE 4.11 This Figure augments Figure 4.7 with a task representation. The proximity compatibility principle (PCP) states that the display representation should be compatible with the task representation. The Figure also suggests that the physical system influences the task representation, which influences the user's mental model in turn.

3. THE THIRD DIMENSION: EGOMOTION, DEPTH, AND DISTANCE

3.1 Direct and Indirect Perception

Much of our previous discussion has focused on two-dimensional (2D) displays. However, there are situations in which a third depth dimension is represented, such that objects in a three-dimensional (3D) scene are represented at various distances from the observer along an axis perpendicular to the plane of the display. These displays are intended to represent three dimensions of Euclidean space, and they will be the focus of the current section. Such displays may be developed for one of two general purposes. First, the three displayed dimensions can represent the three spatial dimensions of physical space, as when a display is constructed to guide the pilot in a flight path, or to plan the trajectory of a robot arm for manipulating hazardous material. Second, the display may use the third (depth) dimension to represent another (non-distance) quantity. Examples of this usage are found in many 3D graphics packages, discussed earlier (see also Chapter 5).

Psychologists have reached broad consensus that there are two qualitatively different systems for perceiving 3D space (DeLucia, 2008). As shown in Table 4.1, these systems have different names, functions, and pathways in the brain (Goodale & Milner, 2005; Patterson, 2007). Importantly for engineering psychologists, they also have different implications for design and multi-tasking (see also Chapter 10).

We describe first a system for **direct perception**, which functions somewhat automatically and is designed for perceiving nearby objects and surfaces as we move through the 3D world, a process called **egomotion**. It is sometimes said to characterize **ambient vision** (Leibowitz, 1988; Previc, 1998, 2002), and its visual receptors are distributed more or less equally all across the visual field (and retina), both in the fovea and periphery. It employs **dorsal visual pathways** leading to the cortex. Its operation in egomotion does not depend heavily upon higher cognitive inference, and so its properties are well represented by the dynamic geometry of the visual image. Because of this anchoring of direct perception in the environment, it is closely associated with **ecological psychology** (Gibson, 1979; Warren, 2004).

In contrast, a system for **indirect perception** is much more dependent on inference and higher-level cognition. This system is useful for more explicit, deliberate judgments of depth and distance of objects, including those objects that are relatively far away from the observer. For instance, this system might be used to judge which of two distant airplanes are closer to a ground observer, or the direction one of the planes is pointing. It makes use of **focal** (usually foveal) **vision**, using **ventral visual pathways**, as opposed to ambient (or peripheral) vision (Previc, 1998, 2000, 2004; Previc & Ercoline, 2007). Because of the use of higher-level cognition, indirect 3D perception imposes a burden on top-down processing and expectancies, in order to make

TABLE 4.1 Two Perceptual Systems

Direct Perception	Indirect Perception
Relatively automatic	Cognitive inference
Egomotion (close to observer)	Object perception (all distances)
Ambient (peripheral) vision	Focal (foveal) vision
Dorsal pathways	Ventral pathways
Ecological	Information processing

depth and distance inferences. This stands in contrast to the relatively automatic processing used for direct perception. Thus, indirect perception places greater demand on attentional resources (see Chapter 10) than direct perception.

When we consider our perception of a 3D environment, both types of perception—direct and indirect—are important. To structure the remainder of this chapter, we will focus first on direct perception and egomotion and its importance for vehicular control. Then we consider the importance of indirect perception and deliberate perceptual judgment for the design of spatial displays.

3.2 Perception of Egomotion: Ambient 3D

As we move through an environment, whether in a plane, an automobile, or on foot, our judgments of the direction and speed with which we are moving depend on information distributed across the visual field, not just in the area of foveal vision (Geisler, 2007; Schaudt, Caufield, & Dyre, 2002). Thus, good drivers who primarily fixate far down the center of the highway are still making effective use of the flow of texture beside the highway as viewed in peripheral vision. As a consequence, engineering psychologists have argued that conventional aircraft navigation instruments (like the attitude display indicator shown in Figure 4.9) are not fully effective for controlling egomotion because they are restricted to foveal vision. Indeed, it has been shown that the pilot's perception of flight information can be augmented by peripheral displays. One example is the **Malcolm horizon display**, which extends a visible horizon all the way across the pilot's field of view using laser projection (Comstock et al., 2003; Malcolm, 1984). Comstock et al. showed that attitude control was much more accurate with the Malcolm display than without.

A second problem with the conventional aircraft instrument panel is that the information necessary for the pilot to obtain a good sense of location and motion is contained in several separate instruments (Figure 4.12), which must then be mentally integrated. One solution to this integration problem is achieved through the development of integrated 3D displays as described briefly in the last chapter. Another solution lies in the design of ecological displays, which capitalize on the visual cues humans naturally use to perceive their motion through the environment—the cues of direct perception that will support egomotion (Bulkley et al., 2009; Gibson, 1979; Larish & Flach, 1990; Warren et al., 2001). Augmented reality displays (see Chapter 5) can provide optical texture to the peripheral scene (Schaudt et al., 2002). In fact, the cockpits of fifth-generation fighter aircraft (such as the F-35 joint strike fighter) make use of such cues and allow the pilot to see sensor imagery "through the floor" using a head-mounted display (*http://en.wikipedia.org/wiki/Lockheed_Martin_F-35_Lightning_II*, 2011).

What information is provided by the external environment as we move through it? Gibson (1979) identified a set of environmental properties that the visual system can detect to assist in control of egomotion. These properties have sometimes been referred to as **optical invariants** because they represent properties of the light rays that reach the eye (or any surface) and have an invariant or unchanging relationship to the location and heading of the observer, whether walking, driving, or flying. It is perhaps useful to think of each invariant as a mathematical function that holds true across various visual environments. Gibson (1979) identified a number of such invariants, and six are described below.

1. *Texture gradient (compression).* The **compression** of a textured surface indicates the relative distances of different parts of the scene from the observer. The *change* in the compression signals a change in altitude or the angle of slant with which the observer is viewing the surface, as is evident when you compare the left and right panels of Figure 4.13.

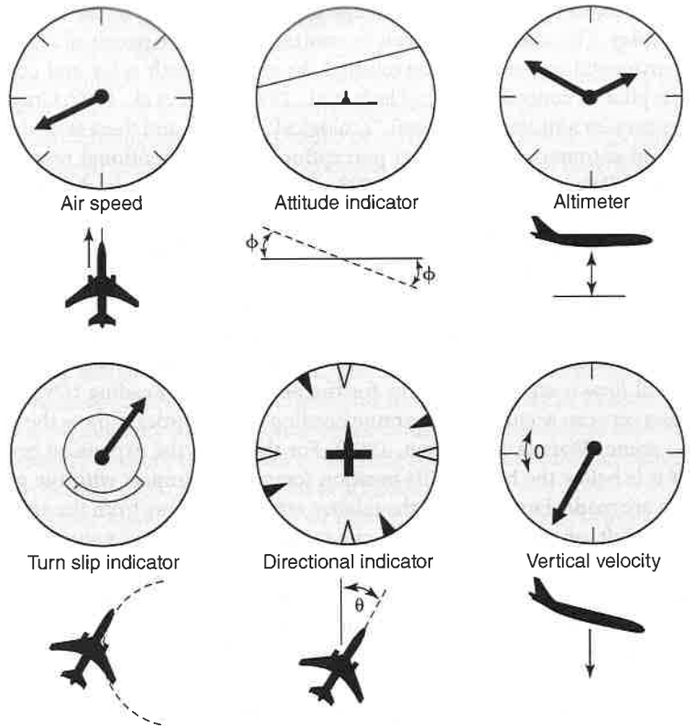

FIGURE 4.12 A traditional flight instrument panel.

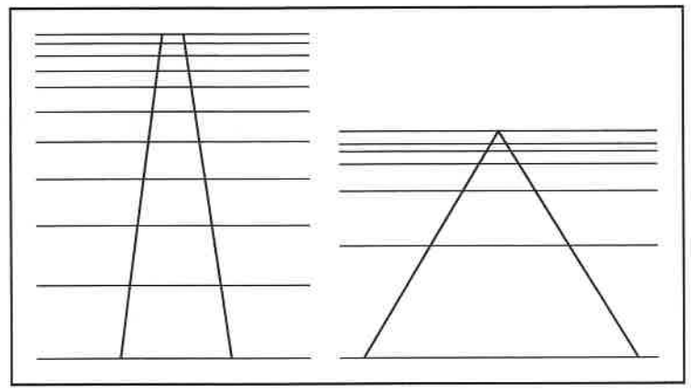

FIGURE 4.13 Splay and compression. Splay is defined by the angle of the two receding lines. Compression is defined by the gradient of separation between the horizontal lines from the front (bottom) to the back (top). On the left, the perception is of being high above the field looking down. On the right, the observer is at low altitude, looking forward. Note how both splay and compression change with altitude.

2. *Splay.* Parallel receding lines signal a change in altitude as given by the angle between the lines—the **splay**. This can again be seen by contrasting the two panels of Figure 4.13.

   Experimental evidence has established the value of both splay and compression in helping the pilot to control altitude (Flach et al., 1992; Flach et al., 1997; Gray et al., 2008). These cues present altitude in a natural, "ecological" fashion, and there is evidence that they are processed automatically by direct perception, leaving attentional resources available for other tasks (Weinstein & Wickens, 1992). Perception of altitude change is particularly important for the airplane pilot to initiate the final stages of landing; pilots make use of the splay of the runway to help determine the altitude (Palmisano et al., 2008).
3. *Optical flow.* **Optical flow** refers to the relative velocity of points across the visual scene (and therefore across the retina) as we move through the world. This velocity is indicated by the arrows in Figure 4.14. The **expansion point** is that place where there is no flow but from which all flow radiates, and it indicates the direction of momentary heading (Warren, 2004).

   Optical flow is an important cue for the perception of heading (Dyre & Anderson, 1997). Observers can accurately determine heading even if optical flow is the only available cue in the scene (Warren & Hannon, 1990). For the pilot, the expansion point is critical because if it is below the horizon, its position forecasts an impact with the ground unless corrections are made. Furthermore, the *relative rate* of flow away from the expansion point, above, below, left, or right, gives a good cue regarding the *slant* of a surface relative to the path of motion. A flow that is of uniform rate on all sides indicates a heading straight into the surface, such as a parachutist would see when descending straight down to the earth. In Figure 4.14, we see that the aircraft is angling into the surface because the optical flow is greater below than above the expansion point. Finally, the *rate* of expansion signals the distance to a surface.

   Greater optical texture density (i.e., more moving points in the scene, more visual detail) generally leads to better control of heading (e.g., Li & Chen, 2010; Warren et al.,

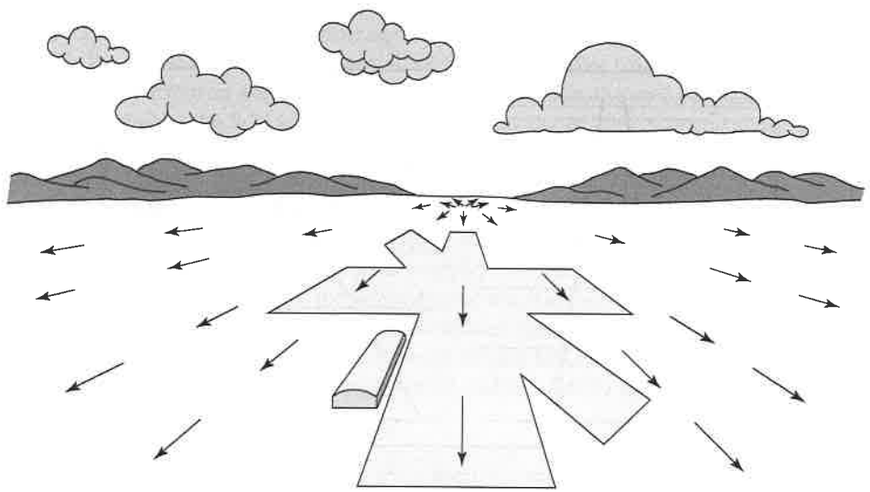

FIGURE 4.14 Optical flow. The arrows indicate the momentary velocity of texture across the visual field that the pilot would perceive on approach to landing.

2001). Thus, if the visual environment is impoverished in terms of optical flow, heading perception will be affected. Kim et al. (2010; see also Palmisano et al. 2008) showed that pilots in a simulator made larger glideslope control errors during landing in night than day conditions, when the terrain texture provides good optical flow.

When landing an aircraft at night over featureless terrain (e.g., when landing over water, darkened areas, or snow), a situation called the **black hole illusion** (Gibb, 2007; Kraft, 1978) can arise in which the pilot thinks he is flying higher than he actually is and descends too quickly, producing a crash or early landing in front of the runway. Through simulation work, Kraft found that, in the absence of the normal textural gradient of the approach terrain (visible on a lighted surface or in daylight, and providing global optic flow), pilots would inappropriately reduce altitude, flying on a dangerously low trajectory that invited ground collision. Several aviation accidents during landing have been directly or indirectly caused by this illusion (Gibb, 2007). One solution lies in the use of virtual imagery on a *head-up display* (or HUD, described more extensively in Chapter 3) to provide the texture: a peripherally located virtual speed indicator using optical flow on the HUD has been shown to be more effective for controlling speed or altitude than conventional cockpit displays (Bulkley et al., 2009; Schaudt et al., 2002).

Consider what happens when we drive in snow (or hail or heavy rain). We have two patterns of optical flow in the environment: one created by our vehicle's motion along the road, and one created by the snow (both by wind and by gravity). The driver's task is to attend to the first optical flow field and ignore the second. However, this is more challenging than it might appear, especially in heavy snow conditions with limited visibility of the roadside. Studies in simulators have shown that drivers tend to drift toward the point of expansion of the snow, rather than that defined by the road and surrounding ground texture. Improved visibility of a simulated roadway has been shown to help drivers maintain course (Dyre & Lew, 2005; Lew et al., 2006). Increased illumination, paint, or signage could be used to produce the same effect on roads subject to heavy snow conditions.

4. *Time-to-contact (tau).* Tau specifies the time remaining until an observer makes contact with an object, assuming that the speed of the observer or the object is constant (DeLucia, 2007; Grosz, Rysdyk, et al., 1995; Lee, 1976). It can be thought of as the rate of change of expansion of an object. Whereas object size and distance are ambiguous (we might be viewing a large object far away or a small object relatively close), the time remaining until contact is unambiguously specified by dynamic information in the visual scene.

   It is clear that observers are sensitive to tau and can make use of it to stop, catch a ball, or take evasive action (Schiff & Oldak, 1990). However, tau is affected by other factors, such as whether the objects are of a familiar size, whether they are partially occluded, or how high the objects are in the visual field (DeLucia, 2004, 2005; DeLucia et al., 2003). These studies suggest that indirect perception can moderate the effects of a directly perceived invariant. We shall return to these ideas when we consider the influence of higher-order cognitive processes on rear-end collisions in the next section.
5. *Global optical flow.* The total *rate* of flow of optical texture past the observer (Larish & Flach, 1990) is determined both by the observer's velocity over the ground and height above the ground. Thus, **global optical flow** will increase as we travel faster and also as we travel closer to the ground.

   Our subjective perception of speed is heavily determined by global optical flow (Dyre, 1997). A potential bias in human perception occurs because perceived speed can appear to increase as height or altitude decreases, even though the actual speed is the same. For example,

we feel as if we are traveling faster in a sports car than in a large sedan or bus, in part because the sports car is closer to the ground. When the Boeing 747 was first introduced, pilots often taxied the aircraft too fast and occasionally damaged the landing gear while turning on or off the runway. The reason for this error, in terms of global optical flow, was simple. The 747 cockpit was about twice as far above the runway as the cockpits in other jets. For the same taxiing speed, the global optical flow was half as fast. Pilots accelerated to obtain a global optical flow that matched their perception of the appropriate taxiing speed established through prior experience. As a result they achieved a true velocity that was unsafe (Owen & Warren, 1987). Similar effects have been found using simulations: observers respond to altitude changes as if they are changes in speed (Wotring et al., 2008). Observers tend to be more sensitive to the global optical flow of the ground when controlling speed, even when they are required to direct their attention elsewhere (e.g., to scan for aircraft above the horizon, Adamic et al., 2010).

6. *Edge rate.* **Edge rate** can be defined as the number of edges or discontinuities that pass across the observer's visual field per unit time. As edge rate increases (texture is finer), the traveler perceives a faster velocity. Global optical flow and edge rate are typically correlated, but edge rate is affected if systematic changes in texture density occur (e.g., if flying and sparse trees change to dense forest), whereas global optical flow is not. Global optical flow and edge rate contribute additively to perception of self-motion (Bennett et al., 2006; Dyre, 1997).

The edge rate cue was exploited by Denton (1980), who was concerned with automobile drivers in Great Britain who approached traffic circles (roundabouts) at an excessive rate of speed. His solution was to decrease the spacing between road markers gradually and continuously as the distance to the roundabout decreased. A driver not slowing down appropriately would see the edge rate as *increasing*. Believing the vehicle to be accelerating, the driver would compensate by imposing a more appropriate degree of braking or slowing. Denton's solution was imposed on the approach to a particularly dangerous roundabout in Scotland. Not only was the average approach speed slower following introduction of the markers, but the rate of fatal accidents was also reduced.

Table 4.2 summarizes our list of optical invariants. As noted earlier, there is increasing evidence that such invariants are most important at smaller distances (less than about 30 m; DeLucia, 2008). If you examine Figure 4.14 carefully, it is evident that points closer to the observer move a greater distance across the retina than far points. At shorter distances, depth information has implications for action, and how we interact with the environment. As already discussed, this has implications for vehicular control; it also has implications for the design of virtual environments (as discussed with regard to the black hole effect). An important implication is that there needs to be sufficient optical texture in the scene to allow detection of the invariants. At longer distances, indirect perception becomes more important for interpreting depth. This is the topic of the next section.

TABLE 4.2 List of optical invariants and what each indicates about egomotion.

Invariant: tells you about
Texture: distance, altitude
Splay: altitude
Optical Flow: heading (slant)
Global Optical Flow: velocity (rate)
Edge Rate: velocity (rate)
Tau: contact

3.3 Judging and Interpreting Depth and Three-Dimensional Structure: Focal 3D

To understand the three-dimensional (3D) structure of space, it is important that we can judge the relative depths or distances accurately. The accurate perception of depth and distance is accomplished through the operation of various 3D perceptual **depth cues**. We will describe each of these cues briefly. Readers wishing more detail about the cues should refer to an introductory perception text such as Goldstein (2010). Some cues are characteristics of the object or world we perceive, and others are properties of our own visual system. We refer to these as **object-centered** and **observer-centered cues** respectively.

3.3.1 OBJECT-CENTERED CUES Object-centered cues are sometimes called **pictorial cues** because they are the kinds of cues that an artist could use in a picture to convey a sense of depth. Figure 4.15 shows a 3D scene that incorporates eight of the following cues:

1. **Linear perspective.** When we see two converging lines we assume that they are two parallel lines receding in depth (the road). This cue is analogous to *splay*.
2. **Occlusion.** When the contours of one object occlude (block) the contours of another, we assume that the occluded object is more distant (on the right, the front building occludes part of the rear building).
3. **Height in the plane (relative height).** We normally view objects from above; when this is the case objects higher in the visual field are farther away (compare the two trucks).
4. **Light and shadow.** When objects are lighted from one direction, they normally have shadows that offer some clues about their orientation, 3D shape, and distance (the buildings and trucks). Although not shown in the figure, lighted surfaces can produce reflectances that indicate the depth of the reflecting object.

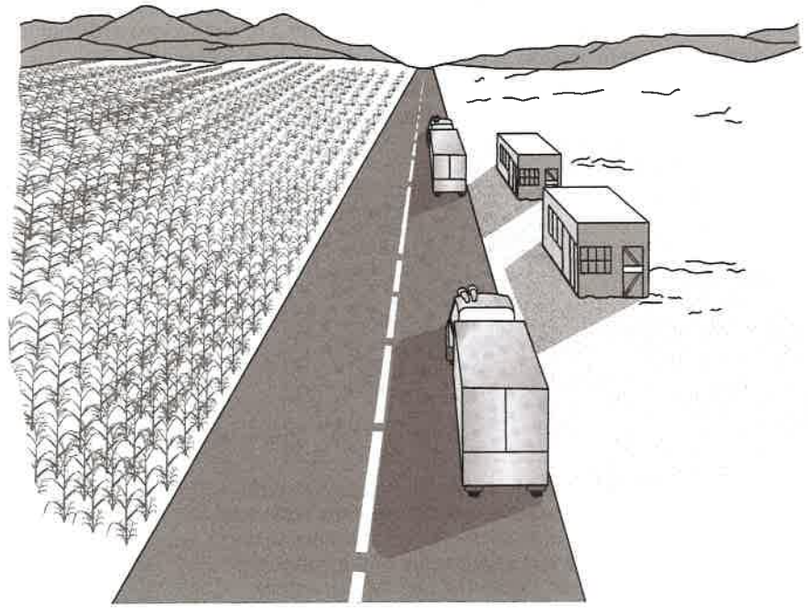

FIGURE 4.15 Contains object-centered cues for depth, as described in the text.

5. **Relative (familiar) size.** If two objects are known to be the same true size, the one subtending a smaller visual angle (smaller area of the retina) is assumed to be farther away (compare the two trucks).
6. **Textural gradients.** As noted when we discussed invariants, the grain on a textured surface grows finer as distance increases (the field on the left and the center line of the road).
7. **Proximity-luminance covariance.** Objects and lines are typically brighter as they are closer to us. The reductions in illumination and intensity with distance therefore signal receding distance (the road lines).
8. **Aerial perspective.** More distant objects often tend to be "hazier" and less clearly defined (the corn field).
9. **Motion parallax.** We use motion information to judge the distances of different objects in the scene. For instance, when we look out a window on a moving train, objects that are closer to us show greater relative motion than those that are more distant. Hence, our perceptual system assumes that distance from us is inversely related to the degree of motion.
10. **Structure through motion.** Motion can be used as a cue to the three-dimensional shape of objects. For example, the cloud of points in Figure 4.16 does not appear to be three-dimensional. Yet if these were points of light on a rotating cylinder, they would show a pattern of motion—slow near the edges, fastest at the center—that leads to an unambiguous interpretation of a rotating three-dimensional cylinder (Braunstein, 1990).

3.3.2 OBSERVER-CENTERED CUES Three sources of information about depth are functions of characteristics of the human visual system.

1. **Binocular disparity (stereopsis).** The images received by the two eyes, located at slightly different points in space, are disparate. Objects at different distances stimulate disparate pairs of points on the retina. The degree of disparity, inversely correlated with object distance, provides a basis for the judgment of distance. Three-dimensional movies and televisions (**stereoscopic displays** discussed in detail Section 3.6) use various artificial methods to present different information to each eye based on this principle.
2. **Convergence.** The "cross-eyed" pattern of the eyes, required to focus on objects as they are brought close to the observer, brings the image onto the detail-sensitive fovea of both eyes.

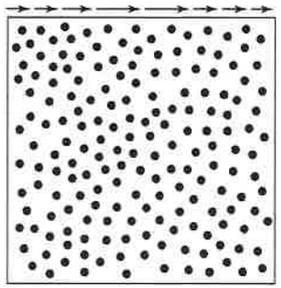

FIGURE 4.16 Potential stimulus for recovery of structure through motion. If the horizontal motion of the dots were proportional to the velocity vectors at the top of the figure, the flat surface would be perceived as a three-dimensional rotating cylinder.

Proprioceptive messages from the eye muscles to the brain indicate the degree of convergence, and therefore the object's distance.

3. **Accommodation.** Like convergence, accommodation is a cue provided to the brain by the eye muscles. The muscles adjust the shape of the lens to bring the image into focus on the retina. The amount of adjustment indicates the approximate distance of the object from the eye.

3.3.3 EFFECT OF DISTANCE ON CUE EFFECTIVENESS The various cues are not all equally effective, and their effectiveness depends on the viewing distance, as shown in Figure 4.17 (Cutting and Vishton, 1995). The figure separates the continuum of depth into three regions: *personal, action,* and *vista space.* Some cues are effective regardless of distance: for example, occlusion and relative size. Other cues tend to be more effective in the different spaces. For example, accommodation and convergence operate only within personal space; within both personal and action space (< 30 m) motion parallax and binocular disparity are important cues for depth. However, as distance is increased, the effectiveness of these cues decrease, and pictorial cues, such as relative size and aerial perspective becomes more important, as illustrated in Figure 4.17.

The range depicted in the figure is based on natural viewing situations. With artificial displays, it is possible to make cues more or less effective at difference distances. For example, stereoscopic displays can artificially represent differences in the distances of objects that are miles away (Allison, Gilliam, & Vecellio, 2009). Furthermore, there are interactions among the cues: while a cue like stereopsis might not play a primary role at large distances, its presence improves visual performance and it appears to validate available monocular cues at large distances (Allison et al., 2009).

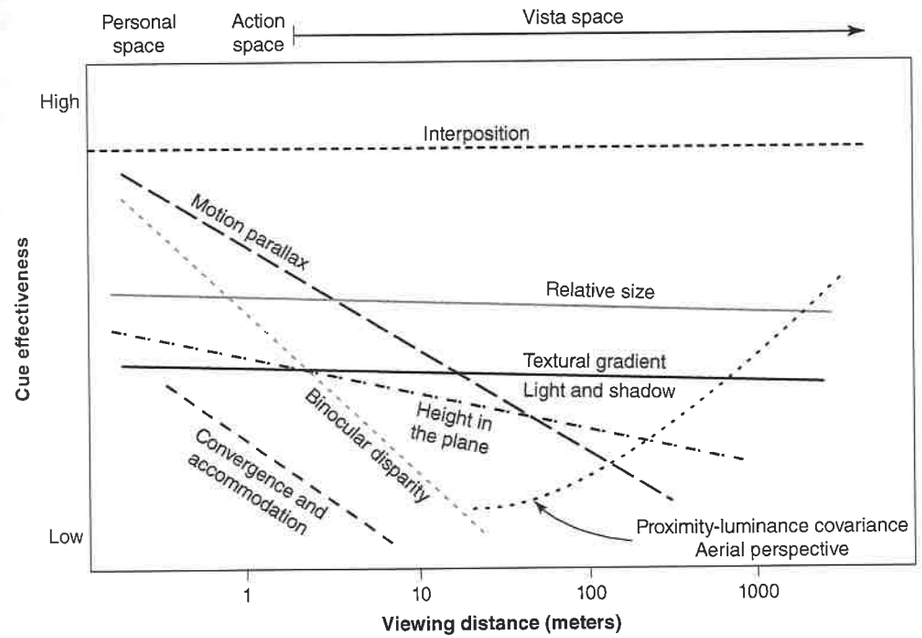

FIGURE 4.17 Effectiveness of various depth cues as a function of distance from the observer.

3.4 Illusions in 3D Viewing

In different ways, Figures 4.15 and 4.17 portray the multiple depth cues that people can use to judge depth and distances in a natural viewing environment. Normally, multiple, redundant cues are available to provide a compelling sense of three dimensionality. In general, the more cues available, the more compelling the sense of depth along the viewing axis (Domini et al., 2011; Wickens, Todd, & Seidler, 1989); however, illusions of depth and distance exist. To understand when depth judgments succeed and fail, it is important to consider how the cues are *integrated* in the brain, an integration that is well explained by the **weighted linear cue model** (WLCM; Bruno & Cutting, 1988; Ichikawa & Saida, 1996; Knill, 2007; Young, Landy, & Maloney, 1993). The model essentially describes the cues as varying in the reliability and precision with which they convey depth information, and through experience with the 3D environment (both short- and long-term; Westheimer, 2011), humans learn to give more weight to more reliable, and hence more *dominant* cues. In this regard, research on depth perception has indicated that three cues in particular tend to be dominant and powerful: *relative motion*, *stereopsis* and *occlusion* (Wickens et al., 1989): they have high weightings in the WLCM.

To illustrate the effects of weighting and **cue dominance**, consider the two objects A and B in Figure 4.18 (top left). Assume that A and B are the same true size. Only a single cue is present, relative size, which suggests that B is farther away (but there is little indication of how *much* farther it is). In Figure 4.18 bottom left, the cue of height in the plane is added, and the sense of depth/distance is more compelling. Now look at 4.18 bottom right. The identical positions and sizes are used as in 4.18 left, but now the near contours of B *occlude* those of A, presenting a clear indication that B is closer. The high dominance of occlusion is demonstrated here (occlusion beats height in plane and relative size).

The importance of the cues in Figure 4.17 in the natural world is found in situations where safety is compromised. This can occur when cues are *insufficient* or *misleading*. We will discuss each of these situations in turn.

When depth cues are missing, there is insufficient perceptual information to provide a compelling sense of depth (we say that the depth scene is *impoverished*). In such cases, like figure 4.18 top,

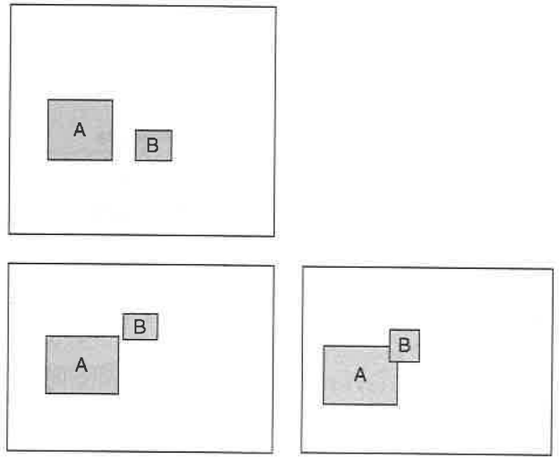

FIGURE 4.18 Illustrating the weighted linear cue model (WLCM). The Figure illustrates the added sense of depth by added cues, and the role of cue dominance by occlusion.

the brain can impose hypotheses on what the depth differences should be, based on past experience and expectancies (Enns & Lleras, 2008; Gregory, 1997; Palmer, 1999). For example, in Figure 4.15 we hypothesize or "assume" that the two trucks in the visual field are the same true size, and therefore the one with the smaller-sized retinal image is farther away. These hypotheses and assumptions are relatively automatic and unconscious. Another example is the black hole illusion, which we described earlier in the context of optical flow (Gibb, 2007; Gillingham & Previc, 1993). When the pilot is flying over dark featureless terrain, there are few cues to the distance of the runway from the cockpit, and the pilot hypothesizes that the aircraft is too high, leading to an aggressive descent.

Even when depth cues are available, they can often be *misleading*. The hypotheses based upon such cues will end up being just plain *wrong*. An example is provided by Eberts and MacMillan's (1985) assessment of why small cars tended to get rear-ended more often on the highway than their larger counterparts. The authors hypothesized, and confirmed with a simulation experiment, the following. The driver behind judges separation, in part on the *relative size* of the vehicle in front, compared to the expected size of the typical vehicle, in order to maintain a safe headway. A smaller car will thus be perceived to be farther away relative to the expected norm; the following car will then inappropriately correct, by pulling too close, and cut the headway to an unsafe margin . . . too close to avoid collision if the small car should suddenly brake. A similar explanation can be offered for why pilots landing at a smaller than expected runway (often a landing strip) will land fast and hard, sometimes overshooting the runway's end (Gillingham, 1993; O'Hare & Roscoe, 1983).

3.5 3D Displays

Understanding 3D perception, and how depth cues combine to provide a compelling sense of depth, is important for the design of 3D displays, especially for those displays that use any and all of the 3D cues in Figure 4.15 to represent depth and distance of real space. The choice of such displays is of course influenced by the principle of pictorial realism (Roscoe, 1968), discussed above. As a result, 3D displays can be very effective formats for representing real space, and we will discuss those success stories first. As discussed earlier, however, the PPR is not the same as naïve realism, which is the commonly held belief that because a 3D display of 3D space is more "realistic," it will always be more effective for spatial tasks (Smallman & Cook, 2010; Smallman & St. John, 2005). People like and want "3D" even when it does not support the most effective task performance. Thus, we will also consider the shortcomings of 3D displays in this section.

3.5.1 3D DISPLAYS OF REAL SPACE One example of such a 3D display is the so called 3D highway in the sky (HITS) display that shows a pilot's commanded route through and actual position within the sky (Figure 4.19; Haskell & Wickens, 1993; Jensen, 1978; Prinzel & Wickens, 2009). The role of relative size and linear perspective in signaling the depth component of the command path is clearly evident in Figure 4.19a, in a way that is missing in the "tri-planar" presentation of the same information in 4.19b. Figure 4.19c presents an example of such a display to be found in emerging versions of corporate aircraft. Several evaluations of this concept have proven it to be more effective than separated tri-planar displays (Prinzel & Wickens, 2008). Within the context of the proximity compatibility principle, the advantage can be seen because flying an aircraft clearly requires integration of motion across all three axes. Hence such an integration task is best supported by the integrated display (Haskell & Wickens, 1993).

In non-aviation domains, 3D displays have also proven superior for tasks in which integration across all three axes of space is required, such as the appreciation of 3D shape, position and trajectory. This would include robotics, industrial or architectural design (Liu, Zhang, & Chaffin, 1997), medical imaging (Hu & Multhaner, 2007), and terrain layout (Hollands, Pavlovic, et al.,

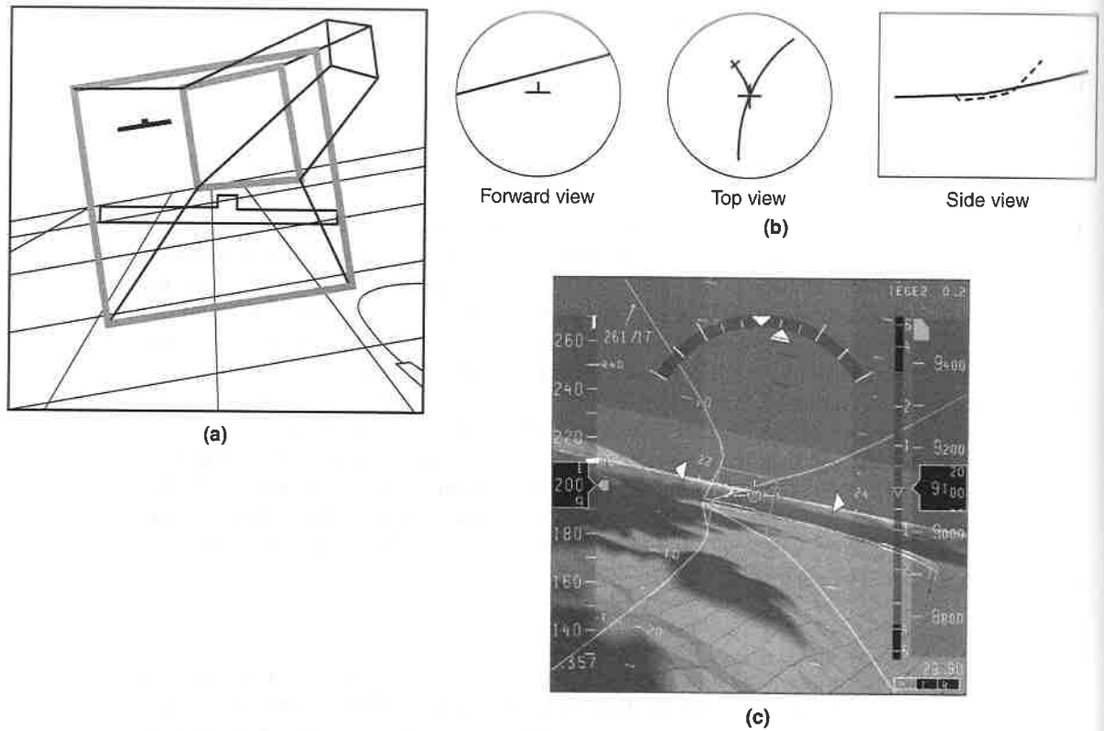

FIGURE 4.19 (a) Highway in the sky (HITS) display. (b) tri-planar representation of the same information. (c) operational HITS display (image courtesy of Erik Theunissen).

2008; St. John, Cowen, et al., 2001; Wickens, Thomas, & Young, 2002). For example, Hu and Multhaner found that resident physicians were better able to determine whether or not to remove a lung tumor using 3D displays of thoracic cavities than they were reading 2D CT images. Tasks requiring shape understanding, such as judging the layout of terrain, or the general shape of 3D objects, are best performed with realistic 3D perspective displays. In Figure 4.20, if you were asked whether you could see point A from point B, you can generally do this better with the realistically shaded, 3D perspective view display (right) than with the plan view topographic map (left) (Hollands et al., 2008; St John et al., 2001).

But 3D displays are not invariably better than their 2D co-planar or tri-planar counterparts (Wickens, 2000a, 2000b). Consider the air traffic displays shown schematically in Figure 4.21. These displays could be used in the air traffic control terminal or as a cockpit display of traffic information (CDTI), which is being introduced into the next generation of aircraft (Alexander, Merwin, & Wickens, 2005; Thomas & Wickens, 2007). Figure 4.21(a) shows a 3D traffic representation. Figure 4.21(b) shows the same information in co-planar form, with the map location of the two planes in the upper panel (X-Y) and the vertical representation of the two in the bottom panel (Z-Y). Here research has shown that the 3D representation of the airspace is inferior for air traffic controllers (May, Campbell, & Wickens, 1996; Wickens, Miller, & Tham, 1996), and either inferior (Wickens, Liang, et al., 1996) or no better (Alexander, Wickens, & Merwin, 2005; Thomas & Wickens, 2007) for pilots. The experimental tasks required controllers or pilots to make judgments

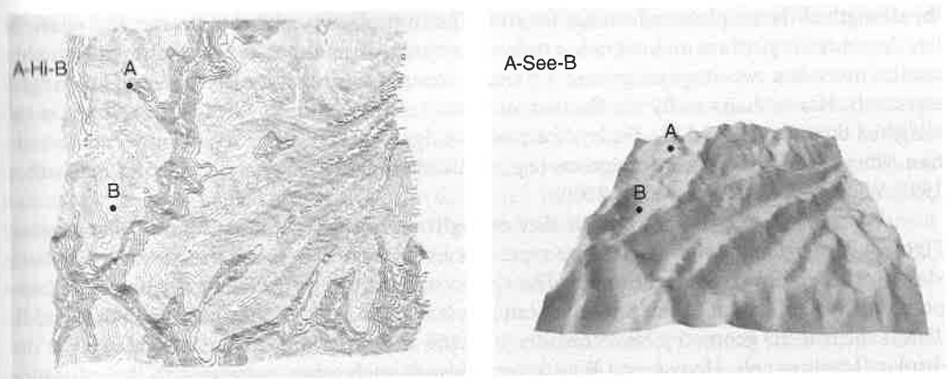

FIGURE 4.20 2D topographic map and a 3D perspective representation of the same terrain. *Source:* 2012 Her Majesty the Queen in Right of Canada, as represented by the Minister of National Defense.

of the proximity or collision risk of aircraft pairs. Such inferiority is observed in spite of the fact that: (a) airspace is 3 dimensional, and hence the 3D display conforms to the principle of pictorial realism; and (b) the judgment of collision risk can be thought of as an integration task, and the 3D display clearly integrates all three dimensional values into a single location in space.

From Figure 4.21, the reason for the inferiority of the 3D ATC display is obvious. The position of the two aircraft is inherently *ambiguous* given that the three spatial dimensions have been collapsed onto a 2D viewing surface (McGreevy & Ellis, 1986). In spite of the added complexity of the co-planar display, the ambiguity is eliminated, and it is possible to *precisely* judge the XY distance (above as the crow flies, over the map) as well as the altitude separation below. In addition,

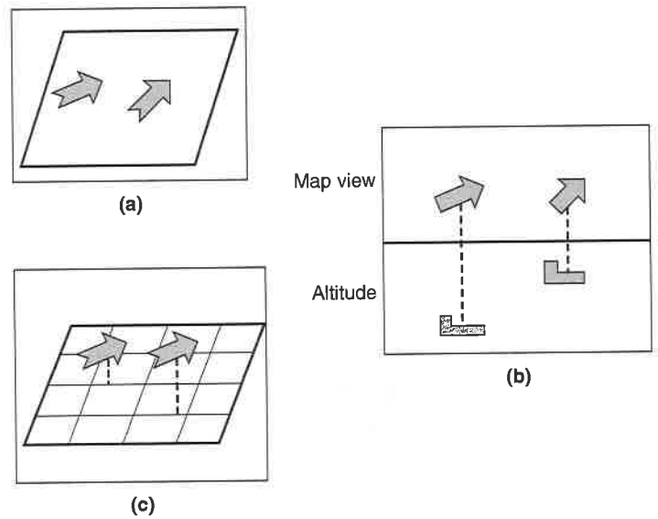

FIGURE 4.21 Three representations of a traffic conflict display, portraying the relative position in 3D space of two possibly conflicting aircraft. (a) 3D, (b) co-planar, and (c) 3D with artificial frameworks.

the strength of the co-planar advantage for air traffic controllers is related to the fact that controllers do not really perform an integration task as they judge separation. Rather, they approach separation more as a two-stage judgment: XY (map) separation, and altitude separation are judged separately. Hence theirs really is a focused attention task. Research in other domains too has established the inferiority of 3D displays for precise judgments along axes requiring focused attention, where the 3D display is ambiguous (e.g., Hollands et al., 1998, 2008; Liu, Zhang, & Chaffin, 1997; Wickens, Thomas, & Young, 2000).

We will unpack the concept of **line of sight ambiguity** (LOS ambiguity) here, using Figure 4.22. At the top of the figure, we represent a volume of space, and the observer's eyeball, viewing this volume from right to left. The space contains three different letter-objects, all approximately equidistant from each other, but A is farther away from the observer than C and B. This is the true 3D geometry. Now consider what the observer would actually see looking at the display (lower panel). Here A and B look very close to each other, compared to their distance from C, a clear departure from the 3D reality. Now suppose the viewer uses the cue of relative size, assuming the letters to be the same true size. Then, seeing the slightly smaller A compared to B, the viewer might realize that A is indeed farther away along the depth or distance axis. But how much farther away? It is impossible to judge, since there are many (indeed, an infinite number of) locations of A along the depth axis and the vertical axis that could produce the same relative position of A and B from the viewer's perspective.

As would be apparent from the WLCM model, discussed above, part of the solution to this LOS ambiguity problem is to provide more depth cues in the image. While this is helpful, when the

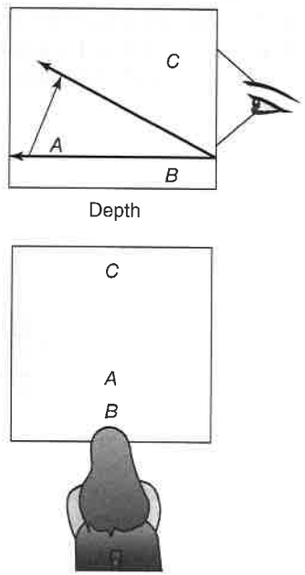

FIGURE 4.22 Top: showing the relative depth along the viewing axis of three objects, A, B and C, as viewed by the observer on the right. Bottom: Depicts the relative position of images on the viewing screen as they would be seen by the observer (represented in the foreground).

3D scene is portrayed in a flat surface, as with a photograph or computer monitor, the benefits of additional depth cues are mitigated somewhat by **flatness cues** (Domini et al., 2011; Young et al., 1993). Here certain features of the viewing environment (e.g., the display frame, reflectance from the screen) signal loud and clear to the observer that this *is* indeed a 2D image. This awareness has a way of perceptually "re-orienting" the perceived depth plane, from one along the line of sight, to one that is progressively more parallel to the viewing screen as depth cues are reduced. This is indicated by the two angled arrows in Figure 4.22 (top). Indeed, if there were no depth cues at all, viewers would perceive all objects to be arrayed vertically on the flat vertical surface. The prominent role of cues to flatness is revealed when those cues are removed. When viewers can no longer see the screen boundaries, or when reflectance is minimized as when viewing the image in a virtual reality simulator, the sense of depth becomes much more compelling, as we describe in the next chapter.

There is also a second cost to 3D displays, closely related, but not identical to LOS ambiguity, and this is compression along the depth axis. Such compression can easily be seen in Figure 4.22 (top). Here the distance between A and B, as viewed on the screen (e.g, in pixels or visual angle) is far less than (is compressed relative to) the distance between B and C. Even when the AB distance is well above threshold, its compression will still degrade the resolution with which differences can be judged (Stelzer & Wickens, 2006) and, for dynamic displays, will reduce the extent to which changes (movement) and changes in changes (rate increases or decreases) can be perceptually resolved. It is of course the low resolution of movement in depth that is responsible for the difficulty in detecting loss of headway in driving, as the car ahead slows down. Similarly, DeLucia and Griswold (2011) showed problems with compression when using multiple camera views in simulated laparoscopic surgery. Performance was poor when their participants used a camera view and the view was parallel to the movement trajectory of the laparoscopic probe.

3.5.2 3D DISPLAYS OF SYNTHETIC SPACE 3D displays can be used to represent conceptual spaces as well as real spaces. In this case, the three spatial dimensions X, Y, and Z are used to represent conceptual variables. Examples would include a 3D scatterplot, a 3D graph like that shown in Figure 4.23, or many of the 3D data visualizations that we will describe in the next chapter. Under such circumstances, while the same limitations of LOS ambiguity and compression apply for focused attention tasks along a single axis, their consequences to performance may not be as serious if precise metric judgments of distance or size are not required. Here the object integration quality of the 3D representation can provide an advantage that outweighs the other costs. For example, when the complex shape of a 3D surface needs to be understood, 3D scatterplot displays have been shown to be superior to separated 2D scatterplots (Kumar & Benbasat, 2004; Wickens, Merwin, & Lin, 1994). However, when precise judgements are required, the costs of the 3D format become evident. For example, if asked to judge the relative heights of two bars in the 3D graph shown in Figure 4.23(a) it is difficult to do this accurately, and the error increases with the distance between the bars in the simulated depth plane (Hollands et al., 2002).

3.5.3 3D DISPLAY SOLUTIONS: ENHANCING DEPTH AND RESOLVING AMBIGUITIES Several remedies to 3D ambiguity can be offered. First, the WLCM suggests that the more depth cues used, the better, and this is clearly supported by research that has varied their number (e.g, Ware & Mitchell, 2008; Sollenberger & Milgram, 1993). Furthermore, given the particularly compelling influence of occlusion, stereopsis, and motion parallax, these should be incorporated whenever possible. Stereo will be discussed in detail in the following section, and motion parallax can be accommodated by allowing the viewer to "rock" or "tilt" the entire displayed volume, much as one might tilt a real 3D transparent volume (like a doll house; Thomas & Wickens, 2007). Flatness cues can be reduced by

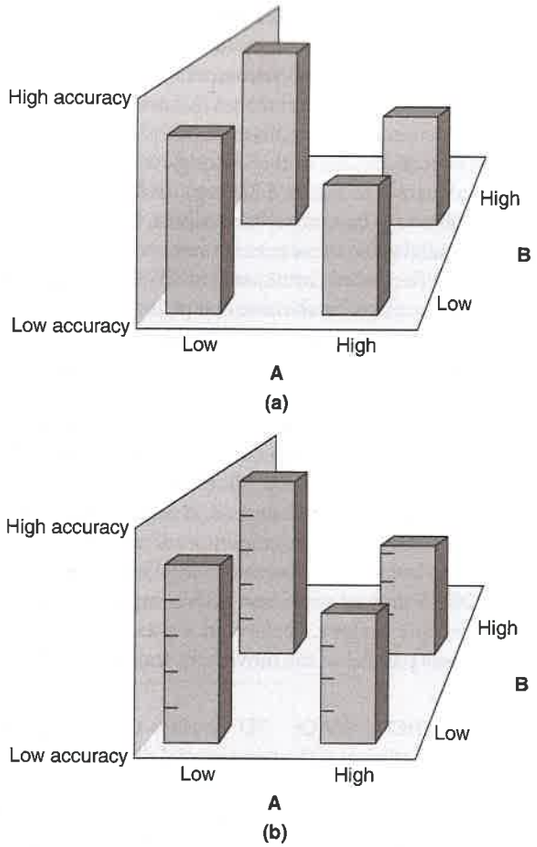

FIGURE 4.23 (a) Perceptual distortions produced by 3D graphics. On the left, the two bars are the same height, but the perception of depth makes the more distant bar appear larger. On the right, the rear bar is smaller than the close bar, but perspective makes them appear the same. Measure the bars to make these comparisons. (b) The same bars are shown with tick marks added. It is now clearer that the two bars on the left are the same size, and on the right, that the bar in front is in fact larger than the bar in the rear.

dimming ambient lighting (eliminating reflection off the display surface), making the display frame less visible, or using immersive VR technology, as described in the next chapter.

Second, artificial frameworks can be added. The tickmarks placed on the bars of Figure 4.23(b) provide a framework that helps judgments of extent (height, in this case). Also, any framework that highlights how differences vary precisely along the 3D orthogonal axes of a volume (lateral longitudinal and vertical) can help. For example, referring now to Figure 4.21c, placing gridlines on the surface and placing the aircraft atop vertical "posts" can help disambiguate their 3D location (Ellis, McGreevy, & Hitchcock, 1987).

Finally, careful *task analysis* is essential. As we discuss in chapter 5, what kinds of cognitive and motor judgments are to be made on the basis of the displayed information? If only holistic judgments or general impressions of space are required (Wickens & Prevett, 1995, call this **global situation awareness**), 3D displays will be superior. But whenever precise judgments along one or more axes are required, co-planar displays should be considered; or the 3D displays should be augmented with an artificial framework. Effective design must accommodate the balance of principles that influence performance of the task required by the user.

3.6 Stereoscopic Displays

As noted above, stereopsis is one of the three dominant cues for 3D depth perception. Indeed many people consider stereo as *the* defining aspect of "3D." We resist this simplistic classification, because motion cues provide a compelling sense of depth when one eye is closed (i.e., without stereo), and indeed monocular viewing can provide a powerful sense of 3D richness from the 10 object-centered cues. Nevertheless, given the importance of the stereo cue, and the technology necessary to generate it artificially, we provide some detail here.

Stereopsis presents slightly different images to the two eyes (Patterson, 2007; Westheimer, 2011). This can be done artificially in a variety of ways. One method is to use glasses with optical shutters that open and close in rapid succession (e.g., at 120 Hz), synchronized with the image shown on the monitor. Another method uses polarized glass so that one lens has horizontally polarized glasses and the other has vertically polarized glass. The display surface depicts two images, each with corresponding polarization. This is the most common method used for 3D movies. The use of different colored lenses works on a similar principle, at the cost of impairing the colors that can be perceived in the scene. Perhaps you have seen 3D bookmarks, cards, or mouse pads in which stereopsis is simulated from a particular viewing angle. These use a *lenticular* printing technology having special lenses that align to control the direction of the light to either the left or right eye. In holographic and volumetric displays, the image is truly 3D, and binocular parallax is preserved in the different directions of light from the display (Patterson, 2007). However, these last methods are challenging to build and require considerable computational power, and as a result are not widely used relative to stereoscopic methods.

As we saw earlier, the amount of disparity can provide a direct, unambiguous cue for depth, and it dominates most other cues with which it is placed in competition. Comparative evaluations generally reveal that stereopsis enhances performance (Getty & Green, 2007; Muhlbach, Bocker, & Prussog, 1995; Sollenberger & Milgram, 1993; Tsirlin et al., 2008; Van Beurden et al., 2009; Ware & Mitchell, 2008; Wickens, Merwin, & Lin, 1994). Stereopsis appears important for the control of limb movement given its high efficacy at short viewing distances. For example, Servos et al. (1992) showed that grasping movements to a target were faster with binocular relative to monocular viewing. In Chapter 3, we talked about the influence of display clutter on visual search and attention; stereopsis can be used as a method for filtering information shown on a display. Kooi (2011) has shown that observers can easily segregate a visual scene on the basis of portrayed depth using stereopsis, which has the net effect of reducing display clutter.

Within the medical community there is great interest in the use of 3D stereoscopic displays for a number of purposes, including diagnosis, preoperative planning, minimally invasive surgery, and medical training (Van Beurden et al., 2009). In general, the advantages of stereo are greatest when visibility is *degraded*, when there is high *scene complexity*, and when there are *few monocular depth cues*. One particular problem for medical imaging systems (e.g., ultrasound, X-rays) is that transparent and translucent surfaces are common and their depiction on a 2D display can be confusing. For example, it can be hard to tell which object is in front (Tsirlin et al., 2008). So for example, Getty and Green (2007) have shown clear stereo advantages for detection rate in breast imaging, reducing both false alarms (false positives) and misses (false negatives).

In preoperative planning, the precise analysis of distances, volumes, and angles is of high importance (Van Beurden et al., 2009). Visualizing multiple intersecting radiation beams to treat a cancerous tumor serves as one example. Again, stereopsis shows clear advantages. For example, determining the optimal path for radiation therapy was performed better using stereoscopic than monoscopic imagery (Hubbold et al., 1997). The advantages of stereopsis for minimally invasive (laparoscopic) surgery appear to be greatest in more complex environments, with more complex tasks, and with inexperienced users (Falk et al., 2001; Votanopoulos et al., 2008). Beyond medical applications, stereoscopic displays will likely be useful for other domains where precise limb positioning and relative position understanding in personal space is necessary.

In summary, stereoscopic displays appear to provide an effective method for increasing the precision of relative position judgments. By reducing ambiguity of depth, they reduce some of the problems observed with 3D displays. However, there are certainly limitations to stereoscopic displays. First, as noted above, they typically require specialized eyewear, which usually produces a drop in the intensity and spatial resolution of an image (McKee et al., 1990; Smallman & Cook, 2010). Second, not all people can accurately use stereoscopic cues. Third, when a richer set of monocular pictorial cues is available (including texture gradient), the advantages of stereopsis can be eliminated (Kim et al., 1987; Ware & Mitchell, 2008). A display designer must balance the added cost of the three-dimensional stereoscopic display against the performance benefits that it provides in a particular task context.

4. SPATIAL AUDIO AND TACTILE DISPLAYS

So far in this chapter we have concentrated on the use of visual displays to depict spatial information. Perhaps this is not surprising, for as we will see when we discuss mental resources in Chapter 11, there is a natural mapping between the visual and the spatial. However, it is certainly possible to use auditory modality to communicate spatial information. An everyday example is the use of stereo headphones, where one musical instrument is placed in the left channel, and another in the right. Tactile displays also have an inherent spatial component. In this section, we briefly address the use of 3D spatial audio technology and tactile displays.

In Chapter 10 we will discuss the role of auditory displays in presenting the operator with information through an alternative channel in order to mitigate the effects of excessive visual workload. Recent advances in computing technology—most notably in the form of **head-related transfer function** filtering techniques—allow sounds to be presented to the listener via everyday stereo headphones that seem to originate from a specific location in 3D space. Under normal listening conditions, we estimate the spatial location of a sound using cues derived from a single ear (**monaural cues**) and by comparing cues received at both ears (**binaural cues**). Similar to the combination of visual depth cues, the monaural and binaural cues are used in combination to determine the location of a sound. If we consider the simple case of the horizontal plane, the auditory system can use differences in both the intensity and timing of the sound as it arrives at each ear. So a sound wave approaching from the left side will reach the left ear earlier, and have greater amplitude (will sound louder), than when it reaches the right ear. So this is a binaural cue. On the vertical plane monaural spectral cues determined by the shape of the pinna are used (Bremen, van Wanrooij, & Van Opstal, 2010). The precise vertical location of a sound is more difficult to determine; although it is mediated by the acoustic context of the sound (Getzmann, 2003). It is through the use of such cues in combination that consumer products with 3D audio technology can reproduce the 3D aspects of the auditory environment, and 3D auditory alerting systems are able to project a sound to a specific location in space, even when the listener is wearing traditional stereo headphones.

The application of 3D audio technologies to aviation has met with considerable success in terms of enhancing performance and reducing workload on a range of tasks, such as target detection and acquisition (Nelson, Bolia, & Tripp, 2001). For example, response times to Traffic Advisory Warning alerts are reduced by 25 percent when 3D audio cues are available (Simpson Brungart et al., 2004). We have a natural tendency to attend visually to loud and distinct sounds, a phenomenon known as the **orientation reflex** (Perrott, Saberi, Brown, & Strybel, 1990), leading to significant decreases in visual search times and improvements in head movement efficiency and effective search area. 3D auditory displays can take advantage of this reflex. Such alerting effects are robust for both static and moving targets and require relatively short training sessions (McIntire, Havig, et al., 2010), are resistant to the effects of sustained high accelerative (gravitational, or G) forces (Nelson, Bolia, & Tripp, 2001), and can also improve the intelligibility of the audio messages themselves (Carlander, Kindström, & Eriksson, 2005). Spatial audio cues can be used to improve the speed of visual search (Pavlovic, Keillor et al., 2009). The location of the auditory cue has to be precise, especially for targets located on the horizontal plane. Even four degrees of error between the target and the sound cue leads to significantly longer search times (Bertolotti & Strybel, 2011).

One advantage to spatial audio is that it more resistant to cognitive load than spoken language. Klatzky, Morrison et al. (2006) guided blindfolded participants along virtual paths. Information was provided to the participant about the azimuth direction of the next waypoint, either using virtual sound or spatial language. At the same time, the participants had to perform a cognitive task (an *N*-back task, to be described in Chapter 7). This task generated a cognitive load for the participants as they tried to navigate between waypoints using the cues. Participants showed better performance while navigating with virtual sound than with spatial language.

Over the last decade **tactile displays** have been developed to present spatial information to operators using tactile actuators. Tactile displays can help direct visual spatial attention, and enhance spatial awareness under degraded visual conditions (Hale, Stanney, & Malone, 2009). Like 3D auditory displays, tactile displays capitalize on the orientation reflex. Tactile displays can reduce spatial disorientation in aviation environments when visual and vestibular cues are missing or misleading (McGrath, Estrada et al., 2004). Tactile displays have also been shown to improve obstacle avoidance (Lam, Mulder, & van Paassen, 2007), facilitate target acquisition for unmanned aerial vehicle operators (Gunn et al., 2005), provide drift information to helicopter pilots during hover (van Veen & van Erp, 2003), and facilitate aircraft upset recovery (Wickens, Small et al., 2008). Like 3D audio, tactile displays are also resistant to the effects of sustained high G forces (van Erp et al., 2007).

The integration of tactile displays with existing visual and auditory displays presents a number of challenges to the designer. One decision relates specifically to whether the tactile cue should provide *status* information (such as the location of an obstacle) or *command* information (tell the operator to avoid the obstacle). Salzer Oran-Gilad et al. (2011) found that for tactile displays used in the cockpit, command displays were preferred over status displays. A related topic (discussed in Chapter 2 in the context of information theory and in Chapter 6 in the context of communications) is the use of *redundancy* to improve performance. Many studies have shown a benefit from simultaneous presentation of the same information through different modalities (for a review, see Wickens, Prinett, et al., 2011). We will revisit many of these topics when we discuss communications in Chapter 6.

In summary, we can see that auditory and tactile displays offer useful methods for presenting spatial information to an operator, if well coordinated with available visual information.

5. TRANSITION

This chapter has described issues related to the design of spatial or analog displays. We began with a discussion of graphs and noted several factors that can make a graph more effective. We then examined graphical displays such as meters and dials, and emphasized the concept of compatibility between the display and the cognitive domain. Then, after introducing two types of perception (direct and indirect) we considered how each contributes to our understanding of 3D space. First, we considered characteristics of a three dimensional environment that provide information about egomotion and how this guides navigation. Then we examined how we deliberately judge and interpret depth and three dimensional structure and discussed how 3D displays might best be designed to effectively convey information. Finally, we briefly considered spatial displays that use other sensory modalities. In the next chapter we will focus on interactive displays that are also spatial, so that chapter forms a natural continuation of many of the topics discussed here. In particular, we build upon and elaborate the discussion of 3D displays. We will address similar topics when we discuss spatial working memory in Chapter 7, and the compatibility between a display and working memory and response in Chapters 7 and 9, respectively. However, as we are well aware, spatial information plays only a partial role in our interactions with other systems, including people. In Chapter 6 we will discuss the complementary role of verbal and linguistic information in such interaction.

Key Terms

Accommodation *111*
Aerial perspective *110*
ambient vision *103*
binaural cues *120*
Binocular disparity (stereopsis) *110*
black hole illusion *107*
brightness *97*
color *96*
color hue *96*
color saturation *97*
compression *104*
Convergence *110*
cue dominance *112*
data-ink ratio *92*
depth cues *109*
direct perception *103*
display compatibility *94*
dorsal visual pathways *103*
ecological compatibility *94*
ecological interfaces *100*
ecological psychology *103*
Edge rate *108*
egomotion *103*
expansion point *106*
flatness cues *117*
focal vision *103*
frequency separated display *98*
global optical flow *107*
global situation awareness *119*
head-related transfer function *120*
Height in the plane (relative height) *109*
hybrid display *97*
indirect perception *103*
inside-out display *98*
Light and shadow *109*
line of sight ambiguity *116*
Linear perspective *109*
magnitude estimation *90*
Malcolm horizon display *104*
mental model *94*
mental operations *88*
meta-analysis *86*
monaural cues *120*
Motion parallax *110*
moving-pointer display *97*
moving-scale display *97*
naïve realism *96*
object-centered cues *109*
observer-centered cues *109*
Occlusion *109*
optical flow *106*
optical invariants *104*
orientation reflex *121*
outside-in display *98*
perceptual continua *90*
pictorial cues *109*
Poggendorf illusion *88*
population stereotype *97*
principle of pictorial realism *95*
principle of the moving part *97*
proximity compatibility principle *86*
Proximity-luminance covariance *110*
Relative (familiar) size *110*
relative judgment or comparison *97*
response compression *90*
response expansion *90*
splay *106*
stereoscopic display *110*
Stevens' law *90*
Structure through motion *110*
tactile displays *121*
tethered display *99*
Textural gradients *110*
ventral visual pathways *103*
visual momentum *93*
weighted linear cue model *112*
work domain analysis *100*

5 SPATIAL COGNITION, NAVIGATION, AND MANUAL CONTROL

The mountain hiker had summited the peak on a beautiful morning and now left the descending ridge to plunge into the wooded valley below, leading to his destination at the distant roadway. The noonday sun gave him a clear orientation along his northbound course. By 1 PM, he had descended below timberline, the sun was now hidden by low clouds, and his GPS unexpectedly gave out. With no compass for a backup, he consulted his guidebook, which indicated that he should take a right turn before the creek drainage. But where was the creek? In a break in the trees he looked upward to find the ridge from which he had descended, but the mountain was now obscured in clouds. He could not match the dim silhouette of the mountain peak with the many humps shown in his map in the guidebook. He thrashed through the trees, came at last to a dirt road, and decided to follow it down. But in the level forest in which he now found himself, which way was "down?"

Much of the material in the previous chapter addressed analog or spatial displays, which are useful for showing continuous differences, such as the slope of a line on a graph, or the position of a pointer on a display. The current chapter also considers issues of continuous representation of spatial information, but does so in the context of location in and *movement* through space (Shah & Miyake, 2005; Taylor, Brunye, & Taylor, 2008). Such movement may be direct, as when walking through a building, along a wooded trail, or hiking a mountain like our lost climber. This movement may also be indirect, as when controlling a bicycle, car, or even controlling a "virtual viewpoint" in virtual reality.

Whether direct or indirect, the movement typically requires some or all of the four primary stages of information processing:

1. A scene or a map must be *perceived* and *attended* in order to find one's current location and goals;
2. The space in which one is traveling must often be *understood*, a process heavily dependent on spatial working memory (Chapter 7). For example: "From what I see, which way is north?" or "Where is the nearest exit?";
3. A direction is *chosen* to meet some task-specific goals, a *choice* that is often based upon the spatial awareness represented in the second stage;
4. The choice is executed through locomotion, either via a simple automated natural method (e.g., walking) or one that may manifest considerable complexity (controlling a large aircraft or submarine in 3D space).

Within this context, the sections of this chapter deal with several related concepts. We begin by describing the cognitive representation of space and in particular, the importance of the **frames-of-reference** concept in spatial thinking (Wickens, 1999; Wickens, Vincow, & Yeh, 2005). In this context, we describe a few important categories of tasks that depend upon this spatial representation. We address human factors tools designed to support these spatial tasks, focusing on