CHAPTER 1

Introduction

Perceiving the world of real objects seems so easy that it is difficult to grasp just how complicated it is. Not only do we need to construct the objects quickly, the objects keep changing even though we think of them as having a consistent, independent existence (Feldman, 2003). Yet, we usually get it right, there are few failures. We can perceive a tree in a blinding snowstorm, a deer bounding across a tree line, dodge a snowball, catch a baseball, detect the crack of a branch breaking in a strong windstorm amidst the rustling of trees, predict the sounds of a dripping faucet, or track a street musician strolling down the road. In all cases, the sensations must be split into that part that gives information about real objects that may change in shape, sound, timing, or location and that part that gives information about the random or non-predictable parts of the background. The object becomes "in front of" the background.

The light energy at the eyes, the sound energy at the ears, and the pressure sensations on hands are neutral. Moreover, the light energy at each eye is two-dimensional due to the "flat screen" structure of the retina, the sound energy at each ear has only a weak spatial component, and the pressure sensations must be integrated to yield the surfaces of objects. For all three senses, the energy must be interpreted to give the properties of the three-dimensional objects and events in the world.

Perception is not passive. Looking at, listening to, sneaking a peek, eavesdropping, rubbing, fingering, shaking, and grasping are all actions that ultimately yield information about objects. But these acts just give us the energy at the receptor, they do not by themselves specify a particular object. Many different objects could produce those same sensations. It is the creative and intentional act of looking or listening that results in the construction of objects (Handel, 2006). Perceptions are not independent of the perceiver. The sensory receptors of all organisms limit what can be known about the real world. It might be too strong to argue that we can never perceive the true physical world, but those limitations act to mark off the consequential features.

S. Handel, *Perceptual Organization*,
https://doi.org/10.1007/978-3-319-96337-2_1

Perceiving is basically focusing on parts of the environment so that depending on their objectives, expectations, and knowledge, people often end up with different outcomes (Felin, Koenderink, & Krueger, 2017).

1.1 The Aperture and Correspondence Problem

Many constraints limit our ability to perceive objects, whether memory or cognitive limitations or environmental obstacles. Auditory, visual, and tactual sensations are constantly changing, and a visual glimpse, auditory snippet, or brief touch must be interpreted in terms of what preceded it in time and space and what will follow it. This has been termed *the aperture problem* to convey the idea that it is like looking through a slit or hole. While this term is couched in visual terms, it obviously also is true for auditory and tactual events that naturally evolve in time and space. The aperture problem is both the cause and a complement to the *correspondence problem*. The sensations at any one instance cannot unambiguously signify objects or events; consequently, the perceptual systems must integrate sensations across time and space to achieve a stable world of objects. This has been termed the *binding* problem—what sensations go with which objects (Burwick, 2014).

The correspondence problem comes in many guises, but the fundamental issue is whether objects or events that occur at different positions, orientations, shapes, intensities, pitches, rhythms, and so on represent the reoccurrence of the same object or represent a different one. I imagine that the observer constructs a trajectory that can link the sensations that occur at different locations and times. Some of these transformations could be predictable (geometrically) particularly for rigid objects. These trajectories or transformations link objects in different orientations, at different pitches, and at different rhythms.

I was driving home one night and noticed two headlights at the same height coming towards me, and naturally assumed that they were mounted on one car, but then the headlights diverged. My first reaction, it was in the 1960s when the TV show *Candid Camera* was extremely popular, that this was a stunt: it was a rubber car or at least a car that was able to shift the position of its headlights. Actually it was two similar motorcycles moving relative to each other. Given my assumption that it was one car, I expected the lights to change position relative to each other based on the geometric properties of rigid bodies, and groped for an explanation when that expectation failed. It is worth noting that I finally realized that they were two motorcycles by the exhaust sounds. Perceiving almost always involves more than one sense. This will be a constant theme in all chapters.

Other transformations would not be as predictable since instruments and singers do not sound alike at widely different pitches and baby pictures do not undergo predictable change as they become adult pictures. Nonetheless, all of these transformations would bind together those pictures or sounds that stem from one object and segregate those that stem from different ones.

The correspondence problem asks whether the "before and after" visual, auditory, and tactual sensations come from the same object. This correspondence might be easy to determine for a single bouncing ball due to its visual trajectory and impact sounds, but it would be much harder to determine which notes correspond to each violinist in an orchestra from the bowing movements.

What this means is that cognitive and physiological constraints along with the uncertainty of the sensory stimulation require the perceptual system to make use of assumptions based on prior experiences with objects in the world in order to perceive accurately. Otherwise, because the sensations could not always correctly make out objects, the world would be ambiguous (Pizlo, 2001). The transformation of the visual, auditory, or tactual sensations into objects depends on many stages and processes at every physiological level. Interactions transform the sensations at the individual receptors in the eye, ear, or hand into information. Cells in the eye, ear, or hand merely respond to energy at different frequencies and pressures. It is the receptive fields in the visual system based on combinations of retinal cells that respond to edges that simultaneously bound objects and separate them from others, it is receptive fields in the auditory system that respond to frequency glides that are characteristic of the sound of many objects (Hubel & Weisel, 1962), and it is interactions of sensors in the skin and muscle that give rise to shape and texture. Further processing isolates the predictable ways objects change over space and time and that predictability shapes our ability to make the world coherent.

The iOS game app Shadowmatic captures this ambiguity by projecting the shadows of familiar objects as those objects undergo various rotations. It is not easy to identify the objects.

In general terms, the organization of the nervous system is a progression from local to global integration (Park & Friston, 2013). Fundamental is that regions of the brain are specialized for a variety of functions such as language, visual recognition, and motor control. To make this work, inputs from different sensory organs are processed in parallel in distinct regions of the cortex. The basic circuitry is similar and hierarchical across modalities: the input ascends to the lower brain centers onto cortical regions, back to the lower regions, and finally back to the cortex (Frangeul et al., 2016). Individual cells are combined into local hubs with dense interconnections that maximize information transfer and processing. These are initially assembled into higher-level global units in development that bring together neural firings from different modalities and further assembled into cortical hubs as illustrated in Fig. 1.1. These higher levels encompass larger spatial regions of the visual, auditory, and tactual fields and longer temporal intervals for all senses; constant feedback at "higher" levels changes the properties of the "lower" levels. Even though information about individual properties (e.g., size, orientation, and location) might be coded in separate "vertical" neural tracts, those tracts need to be integrated because the interpretation of shape is a function of orientation and vice versa. Brincat,

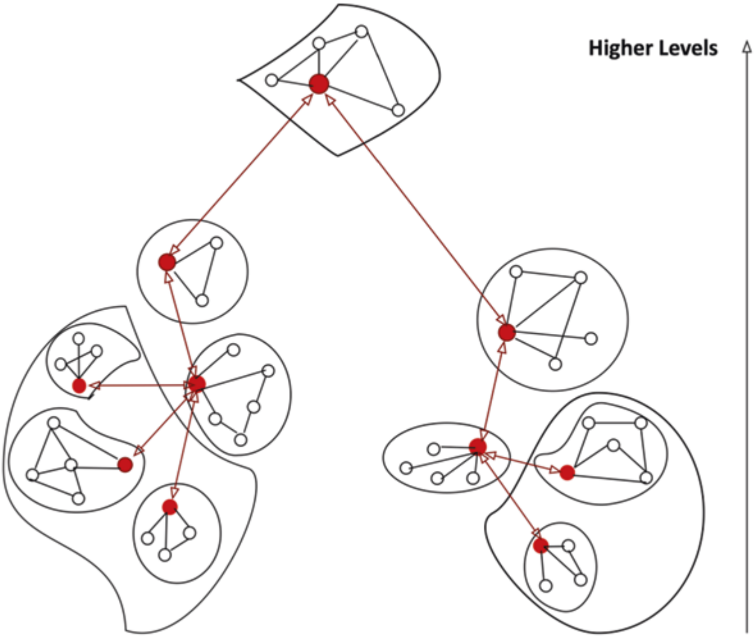

Fig. 1.1 An illustration of the rich club hierarchical modular organization of the brain. Individual nodes (open circles) composed of groups of cells are interconnected anatomically and fire at the same time to individual stimuli. "Rich hubs" (red circles) unify the firings of the individual nodes at a local level. The rich hubs connect at higher levels for specialized functions and ultimately connect at regions of the cortex to create modules for perception and cognition. The connections go from lower to higher modules, but there are feedback connections from the higher levels back to the lower levels depicted by the arrows in both directions. (Adapted from Park and Friston (2013))

Siegel, von Nicolai, and Miller (2018) found that regions in the cortex took on different roles and interacted in unique ways as a function of the task. A hoop changes shape as it rotates vertically. If the levels and tracts were not tightly coupled, it would be impossible to determine if the hoop had rotated or had changed shape (Bassett & Gazzaniga, 2011; van den Heuvel & Sporns, 2013).

To summarize to this point, an accurate construction of the physical world depends on the predictable parts and transformations of objects as well as an integrated nervous system that can abstract those properties as shown in Fig. 1.2. It is this predictability that allows the same object to be identified in different environments. Moreover, visual, auditory, and tactual objects exist in a common extended spatial and temporal framework so that each sense can affect the other. Why you see, hear, or feel what you do is a joint product of the physical stimulation and perceptual processes. In most cases, objects give rise to sensations and information in multiple senses so that cross modal perception is the norm.

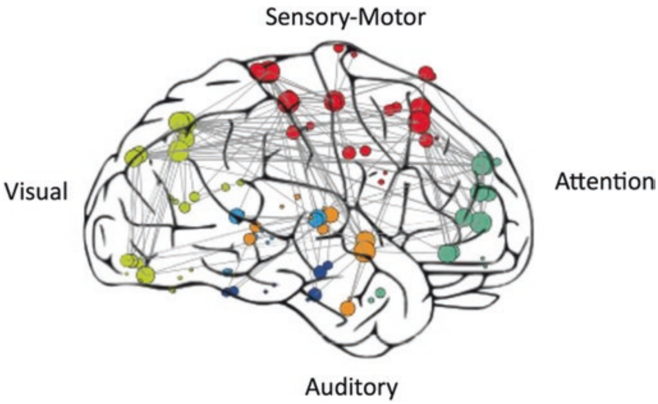

Fig. 1.2 Nodes in localized brain regions underlying specific functions (in different colors) are tightly interconnected; straight lines represent the connections. (Adapted from Bassett and Gazzaniga (2011) and Bullmore and Bassett (2011))

1.2 Similarities Between Perceiving and Business Decision-Making

One way to highlight the assumptions underlying perception is to compare them to assumptions about decision-making in business organizations. While this might seem to be a far-fetched analogy, I think the two have much in common. Both are hierarchically organized with data gathered at the lowest level being successively transformed as it moves up the "ranks." Activities at each level affect all other levels, and purportedly there is a decision-maker at the highest level. Yet there are obvious differences. One is that in organizations the number of workers decreases as one goes up the decision tree, but in perceptual systems the number of neurons in the auditory and visual cortexes is far greater than the number of receptors.

Roberto (2009) lists five myths of business decision-making and the reality underlying each of them and many are comparable to the myths and realities of perceptual decision-making. In business, as in perception, decision-making involves multiple processes and interactions among many levels. The information gathered at the lowest level moves up the system and to a great degree limits the outcomes. Much of the decision-making occurs "off stage," making it difficult to understand how a decision is reached. Current decisions, moreover, are framed by prior outcomes, delimiting which alternatives are considered. As in my headlight example, what you expect to happen can "blind" you to the present.

I have added three more myths (marked by stars) implicit in Roberto's lectures. Businesses in widely divergent fields face the same management problems; success and failure rest on the same solutions. In similar fashion, visual, auditory, and tactual sensations have the same spatial and temporal structures, so that senses must use the same processes to form objects and sources. Moreover, the ability of a business to meet new situations depends on the ability to enlist different parts of the organization to solve the emerging problems interactively. In similar fashion, perceptual systems solve new problems (e.g., hitting a curve ball, playing the banjo) by combining existing perceptual and motor processes in novel ways. Finally, although complex statistical models may yield accurate decision rules, simpler rules of thumb often equal their success rate. In similar fashion, we think the goal of perceiving is to accurately reflect the nature of the physical world, but really the goal of perceiving may be only a successful action, not a perfect representation.

Business decision-making	*Perceptual decision-making*
1. **MYTH**: The chief executive decides. **Reality**: Simultaneous activity at various levels of organization	1. **MYTH**: Single cells decide ("grandmother" cells: one cell represents one person or object) **Reality**: Perhaps as many as 30 million cells are required to represent one image
2. **MYTH**: Decisions are made in the room **Reality**: Real work occurs "off-line"	2. **MYTH**: Percepts determined only in higher cortical centers **Reality**: Processing starting at eye, ear, and hand determines the ultimate percept
3. **MYTH**: Decisions are largely intellectual decisions **Reality**: Decisions are complex social, emotional, and political processes	3. **MYTH**: Percepts are rational, based on sensory information **Reality**: Most percepts emerge from unreflected processes
4. **MYTH**: Managers analyze and then decide **Reality**: Strategic decisions unfold in a non-linear fashion with solutions often occurring before defining the problem or analyzing alternatives	4. **MYTH**: The visual cortex (back of head) combines all inputs and then decides **Reality**: There are many cortical centers (aural, visual, tactual) that transform and constrain the neural signal along the way
5. **MYTH**: Managers decide and then act **Reality**: Decisions evolve over time and involve iterative process of choices and actions	5. **MYTH**: Percepts emerge and resist change **Reality**: Percepts evolve as additional sensory information occurs
*6. **MYTH**: Each business is unique **Reality**: The success and failure of all businesses are based on the same factors	*6. **MYTH**: Vision, audition, and touch are fundamentally different. Vision is spatial, audition is temporal, touch is spatial and temporal **Reality**: Vision, audition, and touch are all based on similar spatial and temporal changes
*7. **MYTH**: Business success depends on the quality of different departments **Reality**: Success depends on continual interactions among departments	*7. **MYTH**: Skills depend on the quality of different brain regions (e.g., musical center, speech center) **Reality**: Skills depend on the brain's ability to reorganize to meet different situations. Perceiving is based on inputs from all senses at once
*8. **MYTH**: Successful decisions depend on exhaustive analyses, possibly based on big data or artificial intelligence **Reality**: Rules of thumb are often equally successful	*8. **MYTH**: The goal of perceiving is to accurately mirror the environment **Reality**: The goal of perceiving is to take the correct action; a perfect rendering of the environment is not necessary

1.3 Summary

As described previously, abstracting things from sensations involves many interdependent perceptual processes. Edges and boundaries arise when abrupt changes in color, brightness, and/or texture break the visual world into discrete objects, when pitch jumps and rhythmic gaps break sounds in discrete sources, and when discontinuous surface features break hard (and squishy) materials into two-dimensional surfaces or three-dimensional objects. Chapter 2 will consider how those visual, auditory, and tactual contours lead to the perception of "figures" in front of "backgrounds" creating a three-dimensional world. But, there are instances in which the ambiguity of the sensation leads to the alternation of plausible but incompatible figures. Even though the most common examples of reversing figures come from simple visual drawings without context, reversing figures also occur in auditory sequences. It is interesting to note here that it is virtually impossible to stop the alternation even if you try to do so by concentrating on one percept. Chapter 3 will show that this alternation is a joint function of cell fatigue and cortical attentional processes. Rhythms also are figures with strong-weak repetitions in time and Chap. 4 will illustrate how auditory, tactual, and visual rhythms are based on levels of accentuation. Chapter 5 will make use of the source-filter model to understand the production and perception of color, timbre, and the process of echolocation. Chapter 6 will summarize these concepts.

References

Bassett, D. S., & Gazzaniga, M. S. (2011). Understanding complexity in the human brain. *Trends in Cognitive Science, 15*(5), 200–209. https://doi.org/10.1016/j.tics.2011.03.006

Brincat, S. L., Siegel, M., von Nicolai, C., & Miller, E. K. (2018). Gradual progression from sensory to task-related processing in cerebral cortex. *Proceceedings of the National Academy of Sciences, 115*(30), E7202–E7211. https://doi.org/10.1073/pnas.1717075115

Bullmore, E. T., & Bassett, D. S. (2011). Brain graphs: Graphical models of the human brain connectome. *Annual Review of Clinical Psychology, 7*, 113–140. https://doi.org/10.1146/annurev-clinpsy-040510-143934

Burwick, T. (2014). The binding problem. *Wiley Interdisciplinary Reviews: Cognitive Science, 5*, 305–315. https://doi.org/10.2002/wcs.1279

Feldman, J. (2003). What is a visual object? *Trends in Cognitive Sciences, 7*(6), 252–256. https://doi.org/10.1016/S1364-6613(03)00111-6

Felin, T., Koenderink, J., & Krueger, J. I. (2017). Rationality, perception and the all-seeing eye. *Psychonomic Bulletin & Review, 24*, 1040–1059. https://doi.org/10.3758/s13423-016-1198-z

Frangeul, L., Pouchelon, G., Telley, L., Lefort, S., Luscher, C., & Jabaudon, D. (2016). A cross-modal genetic framework for the development and plasticity of sensory pathways. *Nature, 538*, 96–98. https://doi.org/10.1038/nature19770

Handel, S. (2006). *Perceptual coherence*. New York, NY: Oxford University Press.

Hubel, D. H., & Weisel, T. N. (1962). Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *Journal of Physiology (London), 160*, 106–154.

Park, H.-J., & Friston, K. (2013). Structural and functional brain networks: From connections to cognition. *Science, 342*(6158), 577–587. https://doi.org/10.1126/science.1238411

Pizlo, Z. (2001). Perception viewed as an inverse problem. *Vision Research, 41*, 3145–3161.

Roberto, M. (2009). *The art of decision making*, *The great courses.* Chantilly, VA: The Teaching Company.

van den Heuvel, M. P., & Sporns, O. (2013). Network hubs in the human brain. *Trends in Cognitive Sciences, 17*(12), 683–696. https://doi.org/10.1016/j.tics.2013.09.012